THE IMPLICATIONS OF UNCERTAINTY IN THE RESULTS OF SIMULATION MODELS AND APPLICATIONS OF ENTROPY MEASURES AS A METHOD FOR UNCERTAINTY QUANTIFICATION IN SIMULATION MODELS

A Dissertation

by

ANNA PAULA GALVÃO SCHEIDEGGER

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee, Amarnath Banerjee
Committee Members, Thomas K. Ferris
Jason Moats
Lewis Ntaimo
Head of Department, Lewis Ntaimo

August 2021

Major Subject: Industrial Engineering

ABSTRACT

This dissertation explores the use of Shannon's entropy and mutual information to quantify uncertainty and to support experimental planning and parameter selection in simulation models. The implications of uncertainty in the results of simulation models are highlighted through an illustrative example where a queue system is modeled using stationary univariate distributions.

In section 2, entropy measures are estimated using histogram-based method with probability density function and discrete empirical distribution. Different number of bins and different normalization methods are investigated. Challenges of working with entropy measures for continuous variables are identified and solutions to these challenges are developed.

In section 3, entropy measures are estimated using kernel-based method, k-nearest neighbors, and fuzzy-histogram-based methods. Different parameters of each method, such as bandwidth, number of k-nearest neighbors, and number of bins, are investigated. This section is an extension of section 2. A different solution to handle the challenges of calculating entropy measures for continuous variables is proposed, which has the advantage of being independent of the choice of the number of bins.

In section 4, entropy measures are applied to investigate the measures' ability to support input parameter selection and experiment planning in simulation models. By using statistical methods, such as regression analysis and contingency analysis, and by comparing the results of the entropy measures with the results from the standard error of

the mean and ANOVA, there is empirical evidence that entropy measures can support the identification of the number of replications that leads to the largest uncertainty and the selection of the most important parameters. With respect to the group of seeds, entropy measures can identify differences among the groups consistently with the standard error of the mean, but not the group of seed that leads to the largest uncertainty.

Overall, the experimental results indicate that entropy measures when estimated using the histogram-based method with discrete empirical distribution appear to be capable to support uncertainty quantification, experimental planning, and parameter selection in simulation models. However, there are still open questions about this topic and directions for further research on this area are articulated at the end of this dissertation.

DEDICATION

To my parents, siblings, family members, and friends who have made this degree possible and who make my life a better experience every day. Especially to my mom, who has given everything to make my path to education possible; and in memory of my godmother who helped shape who I am today.

ACKNOWLEDGEMENTS

First, I would like to thank my committee chair and my committee members for their support and their contributions to my research. I would like to start by expressing my sincere gratitude to my advisor, Dr. Banerjee, who has provided me with guidance, expertise, and support since my first day in College Station. Dr. Banerjee has helped me focus on research, supported my professional endeavors throughout the academic program, and guided me whenever appropriate. He has provided me different research, teaching, and industrial experiences, which had enriched my doctorate degree and had allowed me to become a better professional. By sharing his classroom experiences with me, Dr. Banerjee has taught me how to become a better professor and a better mentor. I would also like to thank my committee members, Dr. Ferris, Dr. Moats, and Dr. Ntaimo, for their consideration, patience, encouragement, availability, and valuable insights into my research whenever I went to discuss my topic with them. Dr. Ferris, Dr. Moats, and Dr. Ntaimo have been an inspiration to me as a professor, mentor, and researcher since the first day I met them at Texas A&M and I am grateful I had the honor to have their support throughout my academic journey.

I would also like to thank the staff of the Industrial and Systems Engineering department. Especially, I would like to thank Erin Roady, Kerra Clement, and Michele Bork. As international students, there is always a handful of paperwork and bureaucracy that we have to go through and having someone that is kind and great on what they do,

but also know how to prioritize things can make a difference in your day. I really appreciate their effort and kindness.

I want to thank the professors of the Industrial and Systems Engineering Department and the Statistics Department of Texas A&M University. Especially, I want to thank Dr. Valdez and Dr. Feldman, for whom I was the teaching assistant for the longest time during my doctorate degree and who were great mentors to myself. I cannot thank them enough for the time, patience, and trust they have deposited on myself and for how much I have learned from them in terms of technical and soft skills.

I want to thank the great friends I made in Aggieland and who were always there for me on the bad and good days. I would like to especially thank: Caro Rodriguez, the first friend I made at Texas A&M, who supported me even virtually throughout COVID times, and whose words of encouragement always moved me forward; Jessie, Caio, Sereninha, and Kobe, the Brazilian family I earned in the United States and who has supported me in unimaginable ways to achieve many of my goals throughout my Ph.D. program, and who I had the honor to share many special moments in Aggieland; Sam, who always patiently listened to my deepest thoughts, who helped me practice my proposal presentation, who has the gift to know the times that I am most in need, and who took care of me and helped me in many different ways during my journey; Nat, who made me laugh with memes during tough times and did not allow me to starve; and many other friends I made in College Station. I would like to also thank my labmates who inspired me through research and who helped me whenever I needed, from teaching assistant hints to

job search hints. Our “worldwide dinner series” was always something I would look for every once in a while.

Finally, I could never have got so far without the much love and support of God, my family, and my dog Maggie. For them, I will always have my deepest gratitude.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Dr. Amarnath Banerjee [advisor], Dr. Thomas K. Ferris, and Dr. Lewis Ntaimo of the Department of Industrial and Systems Engineering and Dr. Jason Moats of the Department of Public Service and Administration.

All the data used in this dissertation was generated through experiments performed in the simulation software Simio® as described in this document. Dr. Banerjee contributed to all sections of this dissertation by providing research ideas, discussing theory, and reviewing the results and the final report. Dr. Ferris, Dr. Ntaimo, and Dr. Moats have provided research insights at the initial phase of the research which contributed to the definition of the research objectives of this dissertation in section 1 and have reviewed the results and the final report.

All work on the dissertation is the result of the student's execution and collaboration with the aforementioned participants.

Funding Sources

Graduate study was supported by the Science Without Borders scholarship of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) under Grant Number [012979/2013-09]. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the CAPES.

This work was also made possible by a fellowship from the Industrial and Systems Engineering Department of Texas A&M University.

NOMENCLATURE

AIC. Akaike information criterion

AMI. Auto-mutual information

ANOVA. Analysis of variance

ApEn. Approximate entropy

ARIMA. Autoregressive integrated moving-average model

AV. Antithetic variates

BME. Bayesian maximum entropy

CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CI. Confidence interval

CIs. Confidence intervals

CMI. Cross-mutual information

CMIFS. Conditional mutual information feature selection

CMIM. Conditional mutual information maximization

CONWIP. Constant work-in-progress

CRE. Cumulative residual entropy

CREM. Council for Regulatory Environmental Modeling

CRN. Common random numbers

CV. Control variates

ECGs. Electrocardiograms

EEGs. Electroencephalograms

ERP. Enterprise resource planning

FD. Freedman and Diaconis

IQR. Interquartile range

IQRt. Information quality ratio

KL. Kullback-Leibler divergence

KNN. K-nearest neighbors

KS. Kolmogorov-Sinai entropy

MADM. Multi-attribute decision making

MAE. Mean absolute error

MaxEnt. Maximum entropy principle

MEGs. Magnetoencephalograms

MI. Mutual information

MIC. Maximal information coefficient

MII. Mutual information index

MinxEnt. Principle of minimum cross-entropy

MISE. Mean integrated squared error

mRMR. Minimum redundancy and maximal relevance

MSE. Mean squared error

NIS. Number of customers in the system

NMI. Normalized mutual information

NMIFS. Normalized mutual information feature selection

NORTA. Normal to Anything

PDF. Probability density function

SAE. Sum of absolute error

SampEn. Sample entropy

SEM. Standard error of the mean

SSE. Sum of squares error

TIS. Time spent in the system

V&V. Verification and validation

VRTs. Variance reduction techniques

VV&UQ. Verification, validation, and uncertainty quantification

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

Simulation models are developed to mimic real systems. Despite the increased details that can be added to simulation models, modelers and researchers acknowledge that a model can seldom precisely reconstruct the real system under investigation due to the system complexity, the large number of variables involved, the dependencies among the variables, the system variability over time, among other factors.

Simulation models are mostly used to represent complex systems and to support critical decisions in terms of economic or social aspects. These models usually show spatial, temporal, and multi-variate dependence that affect the quality and accuracy of their results (Schefzik, Thorarinsdottir, & Gneiting, 2013). Ignoring these dependencies and the consequent uncertainties can lead to over- or under-confidence in the model results (Barton, Nelson, & Xie, 2010).

According to Xie, Nelson, and Barton (2014b), the random input variables are widely modeled as independent univariate distributions in simulation models. However, these input variables may depend on each other or may exhibit a pattern throughout time. Lack of fidelity in the random input models can lead to an inaccurate representation of the system and poor simulation estimates.

Hanson and Hemez (2003) declared that simulation models are also naturally dependent on the modeler's understanding of the system. This idea was reinforced by

Christley et al. (2013) in their paper entitled "'Wrong, but useful': negotiating uncertainty in infectious disease modelling". According to these authors, epidemiological models, as any simulation model, are surrounded by assumptions, approximations, and human influence.

As highlighted by Oberkampf, DeLand, Rutherford, Diegert, and Alvin (2002), a simulation model is always a simplified representation of the reality and any complex system, or even simple ones, contains details that are not represented in the model. Besides, if a system is driven by inputs with randomness, uncertainty will always be present in the model (Biller & Gunes, 2010). Consequently, simulation models are always subject to errors and uncertainty (Marelli & Sudret, 2014).

DeVolder et al. (2002) argued that the more complex the system is, the harder it is to get a precise solution from the model because the uncertainties are also greater. These authors also mentioned that this is somewhat ironic, because models, especially simulation models, are mostly needed to represent complex systems.

To emphasize the trade-off between model complexity and model uncertainty, the Council for Regulatory Environmental Modeling (CREM) showed that the total uncertainty is depicted as a sum of the model framework uncertainty and the data uncertainty. As illustrated in Figure 1, increasingly complex models reduce model uncertainty as more details and, consequently, better understanding are incorporated into the model (Council for Regulatory Environmental Modeling, 2009). On the other hand, more details increase data uncertainty as more input variables and data are required.

Therefore, a trade-off decision must be made between model complexity and uncertainty and there is an appropriate level of complexity that will lead to the minimum total uncertainty.

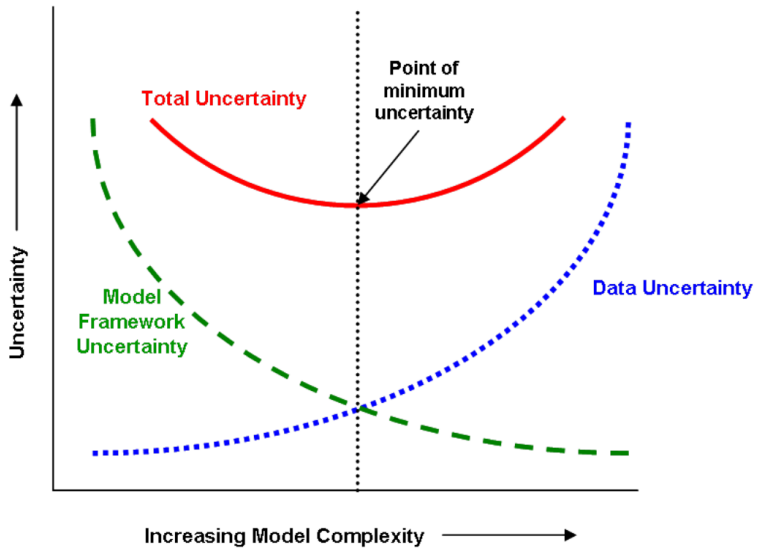

Figure 1. The trade-off between model complexity and uncertainty.
Source: Council for Regulatory Environmental Modeling (2009).

Model uncertainties may not only lead to social or economic losses, but they can also hinder society's trust in using models as a decision support tool (Kitching, Hutber, & Thrusfield, 2005). The usefulness of simulation models depends, then, on controlling their error and uncertainty (Barton, Nelson, & Xie, 2014). However, calibration tools and some other practices may not resolve the issue of model uncertainty. By calibrating some parameters, modelers could be actually ignoring or altering the uncertainty inherent to the

system and, consequently, the uncertainty and errors of the model results could be larger. Roy and Oberkampf (2011) agree that model calibration may not always be the best technique to deal with model inaccuracy. According to these authors, sometimes it may be better to simply account for the mismatch when reporting the results. Barton et al. (2014) added that input-uncertainty error can be aggravated by some practices adopted to control simulation-estimation error. The conclusion to be made is that model uncertainty cannot be eliminated if one wants to correctly represent the system under investigation (Biller & Nelson, 2002).

Although models are inevitably uncertain, the consensus in the academic field is that based on model uncertainty identification and quantification and given the appropriate reflection about this uncertainty, models can still effectively support decision-making. According to Oberkampf et al. (2002), a model with limited, but known applicability, is more useful than a very detailed or complex model with unknown uncertainty. The appropriate reflection on model uncertainty involves informing decision-makers about how uncertain the model results may be, and where, when, and under which conditions the model results are applicable.

Considering the importance of uncertainty for simulation results, DeVolder et al. (2002) claimed that an algorithm or a systematic method was needed to quantify uncertainty in simulation models. These authors believed that to be an effective decision-support tool, simulation models must provide estimates about their level of accuracy and

level of applicability so that decision-makers could determine the appropriate level of confidence to be placed on the results.

Nevertheless, estimating or quantifying uncertainty is not an easy task. Christley et al. (2013) reasoned that only a few uncertainties can be quantified, and even this quantification is most likely uncertain.

With the advances in communication systems, sensors, and computer technology, a large amount of data has become available to the modelers. The increase in data has stimulated modelers to consider larger number of parameters in their studies, even though this data may be sometimes inaccurate or irrelevant.

Two other reasons for modelers to consider more parameters in their studies include: an attempt to better mimic the reality and the limited knowledge about the system under investigation. This last reason implies a larger number of parameters to be tested. Consequently, computer models are usually high-dimensional with respect to the input parameters and, sometimes, even with respect to the responses of interest (outputs).

The increase in the number of parameters in simulation models may lead to a better approximation of the real system, but it can also increase uncertainty. Moreover, an increased number of parameters also means increased resource needs.

To extract meaningful results from a simulation model, a modeler needs to input (accurate) data into the model and run it for an adequate time period, known as run-length, and for an adequate independent number of times, known as replications. Each of these replications is known as an experiment unit. A group of replications with identical settings,

also known as treatment, is called scenarios. The study in which one or more treatments are applied to experimental units is called experiment.

With a large number of parameters, a great deal of experimentation is required (Callao, 2014). The simulation outputs are affected by the numerous input parameters. A larger number of replications is needed to determine whether a parameter has significant influence on the output or how a specific scenario significantly differs from another scenario.

Computer time, as well as modeler time, can be expensive. Besides the costs incurred by increased time requirements, time itself is also a constraint. That is, a modeler's time is limited and the simulation model must provide information in a timely manner in order to be appropriately used for decision-making. In order to minimize the cost and time required for experimentation, either the run-length, or the number of replications, and/or the number of experiments itself must be reduced.

Despite the large number of parameters used in simulation models, the Pareto principle frequently applies (Box & Meyer, 1986a). According to the Pareto principle, a large proportion of changes or effects in the system can be explained by only a small proportion of the input parameters (Box & Meyer, 1986b). The existence of only a few important parameters is referred to as factor sparsity. This means that appropriately selecting the parameters from which to construct the model is of critical importance not only to provide accurate responses, as highlighted by Elizabeth G. Ryan, Drovandi, Thompson, and Pettitt (2014), but also to improve utilization of resources. Eliminating

unimportant parameters at an early stage allows experiments to be run more quickly, with fewer resources, and usually with increased accuracy/reduced uncertainty in the results.

Determining the optimum number of replications to be run is also of critical importance, as running too few replications may not give enough information about the system, but running too many replications may not bring any marginal information to the system as stated by the law of diminishing returns. In this thesis, marginal information refers to the increment in information resulting from a unit replication increment to the total number of replications used in the experiment.

Therefore, to use the resources efficiently, simulation modelers must carefully select the parameters to include in the simulation model, design the experiments that will be run, and plan the configuration of the experiments. The complexity in planning and designing the simulation experiments increases when there is more than one response of interest, because input parameters that are not important for one response of interest, may be critical for another one. A similar challenge occurs when dealing with different estimators: parameters that may affect a measure of dispersion, may not have the same effect on a measure of central tendency and vice versa. Therefore, the choice regarding the appropriate model design and the appropriate experiment design and configuration may also depend on what the modeler is expecting to obtain (Clyde, 2001).

The aforementioned context can be summarized by the following trade-off discussion. Modelers want to incorporate more details (i.e., parameters) into simulation models in order to better approximate the models to reality and, consequently, help

decision-makers better understand the system dynamics. The ultimate goal is to increase the accuracy of the model results. Ideally, modelers would like to include as many parameters as possible. However, knowledge, data, computational power, and time are limited. The available data can also be inaccurate. Therefore, by including more parameters one can, possibly, introduce more uncertainty and errors in the model and, hence, reduce its accuracy. Ignoring the model uncertainty can have negative impacts, such as: over- or under-confidence in the model results, economic losses, accidents, fatalities, and so on. Unfortunately, model uncertainty cannot be eliminated if one wants to correctly represent the system under investigation. Nevertheless, simulation models can still be useful in supporting decision-making, as long as the decision-makers are informed about the uncertainty in the model results. To quantify the simulation model uncertainty, one has to run experiments. Due to limited budget and time, it is important to efficiently plan and design the simulation experiments in order to be as informative as possible, while accounting for uncertainties in the model and for any available prior information while choosing the design of the model. To acknowledge uncertainties in the model results will not only lead to better-informed decisions, but also to a better understanding of the system being modeled.

In this context of increased complexity and uncertainty, limited budget, limited time, and different usability, smart use of the available resources is essential (Clyde, 2001). Clearly, a methodology is required for quantifying uncertainty in simulation models and for understanding the experiment settings that contribute to uncertainty. The

management of uncertainty is essential when the number of replications is limited and the simulation model involves a large number of parameters (Dean & Lewis, 2006). This trade-off discussion is summarized in Figure 2.

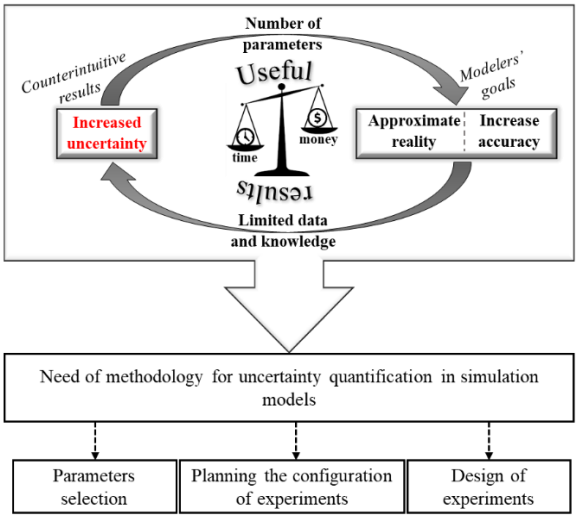

Figure 2. The need for a methodology for uncertainty quantification in simulation models.

The remainder of this dissertation is organized as follows. Section 1.2. discusses simulation experiments that show the relevance of this dissertation topic by exemplifying the implications of uncertainty in the results of a queue simulation model. Section 1.3. provides a brief overview of the topics of this dissertation and their contributions. Section 2 discusses the use of histogram-based entropy measures as a method for uncertainty quantification in simulation models with stationary univariate distributions. Kernel, k-nearest neighbors, and fuzzy-histogram-based entropy measures are discussed in Section

3 as a method for uncertainty quantification in simulation models. In Section 4 different applications of entropy measures in combination with other methods such as regression analysis and Tukey-Kramer multiple comparison test are presented to understand the input parameters and experiment settings (i.e., number of replications and seed) that most contribute to uncertainty in simulation models. Finally, conclusions and future research directions are provided in Section 5.

1.2. Relevance of the Topic: Implications of Uncertainty in the Results of a Simulation Model

As mentioned earlier, there is a trade-off between model complexity and model uncertainty. In general, the more complex the simulation model is, the larger the number of replications one will have to run in order to be able to reach statistically significant conclusions. With the increase in the number of replications, a larger amount of data is obtained about the system. This, in turn, is expected to reduce the uncertainty about the system being investigated. However, one must be careful with this statement. By running more replications, data uncertainty is in fact reduced. Nevertheless, the model uncertainty remains the same and so does the confidence level in making statements about the output of interest. As it is well known, the confidence interval (CI) either contains or does not contain the point estimate of interest. Simply running more replications does not affect the confidence level of the conclusions obtained from the simulation results. This means that by running more replications and, consequently, consuming more resources, one may miss the true estimate about the system, while a scenario with fewer replications and,

consequently, with larger data uncertainty may contain the true estimate. In this context, it is also important to assess if running more replications is economically and computationally attractive. That is, it is important to evaluate whether the additional replications bring substantial marginal information (i.e., increase in information per unit of replication) to justify the additional use of resources. From this discussion, one can see that the implications of uncertainty in the results of simulation models is complex and deserves further attention.

To investigate the implications of uncertainty in the results of simulation models a simple queue model was built using Simio® University Enterprise Edition v 10.165. The model consists of a single source of arrivals, a single queue, and s servers providing the same service. After being served, customers leave the system. Balking and reneging were not considered in the model. Two input parameters were considered in the model, namely inter-arrival time ($1/\lambda$) and service time ($1/\mu$), and two output responses were considered, namely average number of customers in the system (L or NIS) and average time spent in the system (W or TIS).

The queue model was chosen because it has closed-form theoretical solutions for some distributions of the input parameters and it also has good approximations for more general distributions. The notation used in this dissertation follows the $A/S/s$ Kendall's notation, where: A represents the arrival process, S the service time, and s the number of servers. M is used for memoryless distributions and G for more general distributions.

Several scenarios were run with different number of replications, different traffic intensities, different number of servers, different variances, different distributions (M = exponential, and G = normal), different parameter values, and different seeds. The run length of each scenario was 1,825 days, which included 365 days of warm-up. The specified warm-up period was enough for the system to reach steady-state for all the experiments.

The equations used to calculate the exact theoretical values of the two output responses, namely L and W, were taken from Gautam (2012) and are given below.

$M/M/s$ systems have the following exact solution:

$$L = \frac{\lambda}{\mu} + \frac{p_0(\lambda/\mu)^s \lambda}{s!\, s\mu[1-\lambda/(s\mu)]^2} \qquad \text{Equation 1}$$

$$W = \frac{1}{\mu} + \frac{p_0(\lambda/\mu)^s}{s!\, s\mu[1-\lambda/(s\mu)]^2} \qquad \text{Equation 2}$$

where:

$$p_0 = \left[\sum_{n=0}^{s-1}\left\{\frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n\right\} + \frac{(\lambda/\mu)^s}{s!}\frac{1}{1-\lambda/(s\mu)}\right]^{-1} \qquad \text{Equation 3}$$

$M/G/1$ systems have the following exact solution:

$$L = \rho + \frac{\lambda^2}{2}\frac{(\sigma^2 + 1/\mu^2)}{1-\rho} \qquad \text{Equation 4}$$

$$W = \frac{1}{\mu} + \frac{\lambda}{2}\frac{(\sigma^2 + 1/\mu^2)}{1-\rho} \qquad \text{Equation 5}$$

where $\rho = \lambda/(s\mu)$ and σ^2 is the variance of the service time. Equation 5 is known as the Pollaczek-Khintchine equation.

Due to the simplicity of the aforementioned system, one would expect to obtain accurate results in the simulation model even when the simulation model is driven by only a few input parameters (i.e., arrival time and service time). As acknowledged by simulation stakeholders, because the input parameters themselves are usually unknown and contain randomness, adding more parameters to the simulation model does not always reduce uncertainty.

To verify the accuracy of the simulation results, response values computed for L and W using the simulation model were compared with the theoretical true steady-state values of the corresponding responses. A summary of this comparison is presented in Table 1. As discussed in the previous paragraph, a simulation model representing a simple real system like this (queue system) should have low uncertainty in terms of both the extrinsic input-uncertainty and the intrinsic output-uncertainty and, consequently, the theoretical steady-state values of the real system should be within the simulated CI. However, the results in Table 1 indicate that this is not always true. Interesting contradictions are highlighted below.

Table 1. Percentage of scenarios with theoretical values within simulated values.

Model / Number of Replications	Number of Scenarios	Number of theoretical *W* values within the simulated *W* CI	% of scenarios that contain theoretical *W*	Number of theoretical L values within the simulated *L* CI	% of scenarios that contain theoretical *L*	Average of smallest absolute error of *W*	Average of smallest absolute error of *L*
M/M/1	150	120	80.0%	122	81.3%	0.053%	0.072%
10	15	15	100.0%	15	100.0%	N/A	N/A
20	15	15	100.0%	15	100.0%	N/A	N/A
50	15	15	100.0%	15	100.0%	N/A	N/A
100	15	15	100.0%	15	100.0%	N/A	N/A
200	15	15	100.0%	15	100.0%	N/A	N/A
400	15	2	13.3%	2	13.3%	0.092%	0.099%
600	15	7	46.7%	8	53.3%	0.062%	0.073%
800	15	11	73.3%	11	73.3%	0.057%	0.058%
1000	15	11	73.3%	11	73.3%	0.054%	0.056%
1500	15	14	93.3%	15	100.0%	0.002%	N/A
M/M/3	150	132	88.0%	133	88.7%	0.022%	0.023%
10	15	15	100.0%	15	100.0%	N/A	N/A
20	15	15	100.0%	15	100.0%	N/A	N/A
50	15	15	100.0%	15	100.0%	N/A	N/A
100	15	15	100.0%	15	100.0%	N/A	N/A
200	15	15	100.0%	15	100.0%	N/A	N/A
400	15	3	20.0%	3	20.0%	0.028%	0.032%
600	15	11	73.3%	10	66.7%	0.016%	0.014%
800	15	15	100.0%	15	100.0%	N/A	N/A
1000	15	15	100.0%	15	100.0%	N/A	N/A
1500	15	15	100.0%	15	100.0%	N/A	N/A
M/M/10	150	131	87.3%	136	90.7%	0.011%	0.014%
10	15	15	100.0%	15	100.0%	N/A	N/A
20	15	15	100.0%	15	100.0%	N/A	N/A
50	15	15	100.0%	15	100.0%	N/A	N/A
100	15	15	100.0%	15	100.0%	N/A	N/A
200	15	15	100.0%	15	100.0%	N/A	N/A
400	15	2	13.3%	6	40.0%	0.010%	0.018%
600	15	11	73.3%	12	80.0%	0.009%	0.014%
800	15	14	93.3%	14	93.3%	0.014%	0.013%
1000	15	14	93.3%	14	93.3%	0.013%	0.012%
1500	15	15	100.0%	15	100.0%	N/A	N/A
M/G/1	140	15	10.7%	27	19.3%	0.875%	0.939%
10	14	2	14.3%	3	21.4%	0.882%	0.918%
20	14	2	14.3%	3	21.4%	0.864%	0.893%
50	14	2	14.3%	3	21.4%	0.917%	0.973%
100	14	2	14.3%	3	21.4%	0.891%	0.935%
200	14	2	14.3%	3	21.4%	0.898%	0.946%
400	14	1	7.1%	2	14.3%	0.830%	0.867%
600	14	1	7.1%	3	21.4%	0.851%	0.981%
800	14	1	7.1%	3	21.4%	0.863%	1.001%
1000	14	1	7.1%	2	14.3%	0.868%	0.926%
1500	14	1	7.1%	2	14.3%	0.885%	0.951%
Total	**590**	**398**	**67.5%**	**418**	**70.8%**	**0.240%**	**0.262%**

The smallest absolute error of W or L was calculated using Equation 6.

$$Smallest\ Absolute\ Error\ of\ Y = min\left(\begin{matrix}\left|\frac{Y_{theory} - Lower\ Bound\ CI_Y}{Y_{theory}}\right|, \\ \left|\frac{Y_{theory} - Upper\ Bound\ CI_Y}{Y_{theory}}\right|\end{matrix}\right) \quad \text{Equation 6}$$

where:

Y is the variable under investigation (W or L);

Y_{theory} is the theoretical true value of W or L;

$Lower\ Bound\ CI_Y$ is the lower bound of the simulated confidence interval of the variable Y; and,

$Upper\ Bound\ CI_Y$ is the upper bound of the simulated confidence interval of the variable Y.

As the results shown in Table 1 indicate, the simulation model performed well, but it was not 100% accurate in representing the queue system under the different scenarios investigated. About 85.1% of the simulated confidence intervals (CIs) contained the true W value of the $M/M/s$ system and 86.9% of the simulated CIs contained the true L value. However, only 10.7% of the simulated CIs contained the true W value of the $M/G/1$ system and only 19.3% of the simulated CIs contained the true L value. The average absolute errors (including or excluding zeros) of W and L did not exceed 1% in any of the scenarios investigated. The individual absolute error of W and L did not exceed 0.5% in

any of the $M/M/s$ scenarios investigated and it did not exceed 3.5% in any of the $M/G/1$ scenarios investigated.

Interestingly, for $M/M/s$ systems, the highest individual absolute error was always observed for 400 replications, regardless of the configuration. These results are surprising for two reasons. First, $M/M/1$ queue systems have exact theoretical solutions and, hence, the error is not due to the numerical solution. Second, with a higher number of replications, the intrinsic output-uncertainty tends to decrease due to the increase in the sample size. Consequently, a few possible explanations for these contradicting results are: (1) the pseudorandom number generator of the software is not appropriate, (2) the warm-up period is not long enough, (3) the extrinsic input-uncertainty is present and significant in the model, and/or (4) the intrinsic output-uncertainty is not monotonically decreasing with the increase in the number of replications. The first explanation is not adequate because Simio® uses the Mersenne Twister pseudorandom number generator that has an extremely long period and has been extensively tested for uniformity and independence. The second explanation can also be eliminated because the effect of the warm-up period was investigated for all the scenarios and the warm-up period used in the experiments was considered satisfactory in all the scenarios. Moreover, if the warm-up period was the issue, the effect should be even higher for scenarios with identical configuration and a larger number of replications. As one can see from Table 1, the error tends to reduce again with a higher number of replications. Therefore, the third and fourth explanations are the only ones remaining unexplained and indicate the impacts that uncertainty may have on the

simulation results. Figure 3 and Figure 4 show the boxplot of six different experiments, with identical $M/M/1$ configuration but with different numbers of replications for the simulation responses time in the system and number in the system, respectively. Experiments are numbered from the lowest to the highest number of replications, where experiment 376 is the $M/M/1$ system replicated 400 times. As can be seen in Figure 3 and Figure 4, experiment 376 is the experiment with the largest error between the simulation responses and the theoretical responses among the experiments shown.

For $M/G/1$ systems, the largest individual errors were obtained in the scenarios with the largest number of replications. Another interesting observation is that despite the fact that the L response had a larger number of true values contained within the simulated CIs, the highest absolute errors observed referred to this response and not to the W response. This was observed for the majority of the scenarios under investigation.

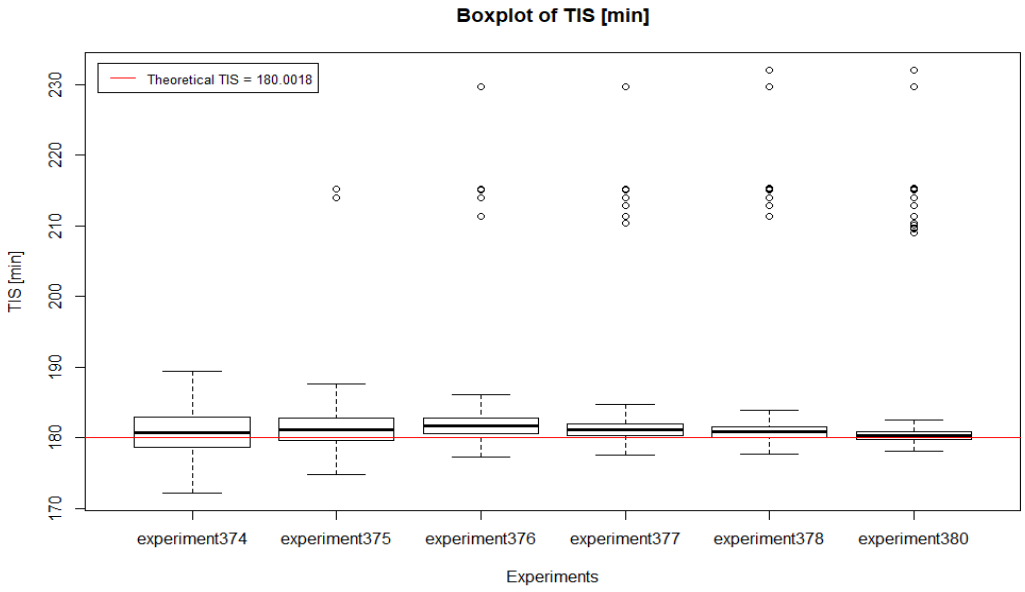

Figure 3. Boxplot of simulation response time in the system.

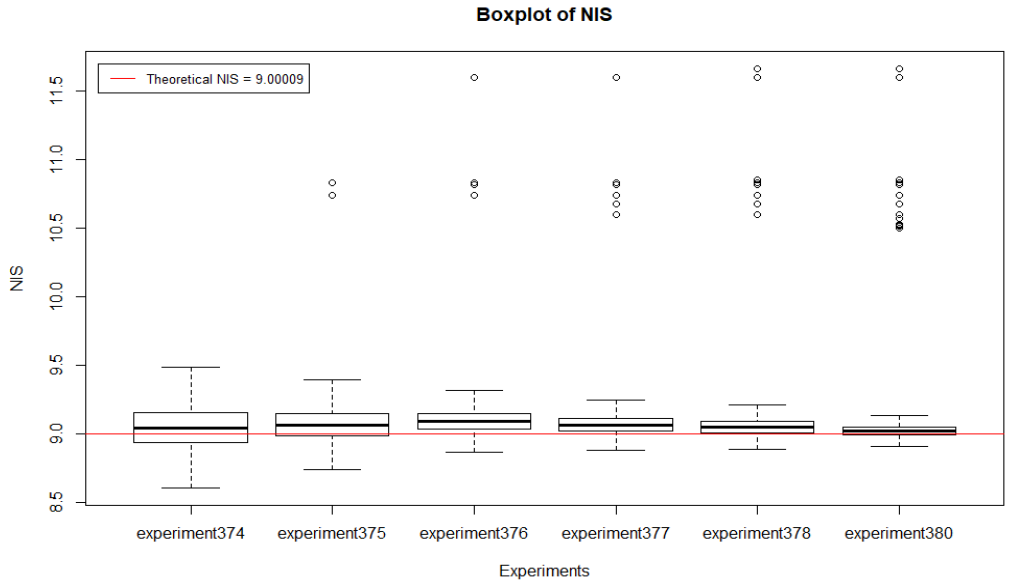

Figure 4. Boxplot of simulation response number in the system.

Although this is a simple example, the situation can be actually worse, because the uncertainty level in more complex systems may be much higher than the ones observed here. For instance, a realistic simulation model usually mimics a complex system that has tens or hundreds of input parameters and at least a handful of responses of interest. In such complex systems, extrinsic input-uncertainty, as well as intrinsic-output uncertainty, are more likely to be higher due to the greater number of assumptions and approximations, higher chances of data measurement errors, greater inherent variability, and so on. Moreover, in more complex systems, there is also a higher probability that the responses of interest will be a correlated nonlinear function of the inputs. Therefore, this illustrative example shows the need for an uncertainty quantification analysis and careful use of the simulation results.

Law (2007) emphasized throughout his book that simulation models are driven by random inputs and, consequently, will produce random output. He argued that appropriate analysis of the output is critical if the results are to be properly interpreted and used. Barton et al. (2014) highlighted that not only corporate profitability may depend on the decisions informed by simulation results, but also human lives. Therefore, simulation modelers should search for systematic ways of improving the model results accuracy and estimating the uncertainty level of the results in order to appropriately inform the simulation stakeholders.

In this context, it is important to reinforce the fact that simulation results contain uncertainty does not mean the results cannot support decision-making or have no use.

What is important is that the decision-makers and simulation stakeholders are adequately informed of these uncertainties and the consequent risks involved in their decisions. As the quote says: “something is better than nothing”. A partially informed decision has higher chances of success than a simple blind guess.

1.3. Research Topics and Contributions

Based on the trade-off between model complexity and uncertainty in simulation models, this dissertation investigates the use of Shannon’s entropy and mutual information as measures of uncertainty quantification in simulation models. The main research question of this work is whether entropy and mutual information measures can quantify the uncertainty and, consequently, the information present in simulation models and help to understand the input parameters and experiment settings that contribute more to uncertainty in simulation models. The work is divided into three topics. The first part addresses the case where the simulation models are driven by stationary univariate distributions and the entropy measures are calculated using the histogram-based method with and without normalization. The second part is an extension of the first part, where kernel-based method, k-nearest neighbors, and fuzzy-histogram method are used as the entropy estimators. Finally, in the third part, the entropy measures are applied to identify the input parameters and the experiment settings, such as number of replications and seed, that contribute more to uncertainty. The contributions of each topic are detailed below.

Under topics one and two, Shannon’s entropy and mutual information, from the information theory field, are used to quantify how uncertain is(are) the simulation

outcome(s) and how much of this uncertainty can be attributed to each of the inputs. The topic of uncertainty quantification has been widely investigated by researchers and according to Barton et al. (2014), because there are robust methods for controlling and measuring simulation-estimation error, one might be tempted to say that the problem of estimating uncertainty has already been solved. Recent studies have concentrated on estimating the input-uncertainty and providing confidence intervals that account for both the simulation-estimation error as well as the uncertainty about the input models. In the case of parametric distributions, input model uncertainty is reduced to input parameter uncertainty (Song, Nelson, & Pegden, 2014). Barton and Schruben (1993) used uniform and bootstrap resampling to estimate simulation output-uncertainty due to the uncertainty in the empirical distribution used as input to drive the simulation. Barton et al. (2010) and Barton et al. (2014) used metamodel-assisted bootstrapping to provide a confidence interval that included the input-uncertainty from independent parametric input models and the simulation-estimation uncertainty. Xie, Nelson, and Barton (2014a) introduced a Bayesian framework to provide credible intervals for the mean that accounted for the input-parameter uncertainty and the simulation-estimation uncertainty. The uncertainty about the input parameters was represented via a posterior distribution conditional on the real-world data and the uncertainty about the mean simulation response via a posterior distribution conditional on a designed simulation experiment. Xie et al. (2014b) used metamodel-assisted bootstrapping to provide a confidence interval that included the input-

uncertainty from dependent Normal to Anything (NORTA) input models and the simulation-estimation uncertainty.

In contrast, this dissertation investigates entropy measures as a method to estimate the total uncertainty in simulation models and to quantify the amount of uncertainty in the simulation response that can be attributed to each one of the inputs. The goal is to provide a method for estimating the uncertainty on simulation responses that requires no additional simulation effort and that is simple enough to be understood and implemented by the majority of simulation stakeholders.

According to Song et al. (2014), decomposing the uncertainty, also known as measures of contribution, is difficult and it is an area of active research. Song et al. (2014) discussed what is called sample-size sensitivity. This technique quantifies how much the estimator variance would be reduced by observing one more real-world sample data from an input process x, given that one already has m observations. The input distributions with the largest sensitivities would be targeted for more real-world data. However, Song et al. (2014) highlighted that sample-size sensitivities are only local-gradients and, hence, not ideal for budget allocation. Ankenman and Nelson (2012) presented a bootstrap method to assess the input uncertainty relative to the simulation sampling variability and they proposed a follow-up experiment using sequential bifurcation to identify the largest sources of input uncertainty. Song and Nelson (2013) build on Ankeman and Nelson's (2012) work. Song and Nelson (2013) used bootstrap and variance decomposition to estimate the relative contributions of each input model and to identify the input data

sources from which collecting more data would lead to the most reduction in input uncertainty. However, in contrast to the work of Ankenman and Nelson (2012), Song and Nelson (2013) provided a new follow-up analysis in which no additional simulation experiments were required. Table 2 lists some of the important works in the field of uncertainty quantification and the methods used by them.

Table 2. Some of the important work in the field of simulation uncertainty quantification.

Paper	**Method**	**Assumptions**
Barton and Schruben (1993)	Uniform and bootstrap resampling	Independent univariate empirical distributions.
Barton et al. (2010)	Meta-model-assisted bootstrapping (stochastic kriging meta-model), direct bootstrap, Bayesian bootstrap, conditional confidence interval, and Bayesian credible interval.	Independent univariate parametric distributions with known families (exponential) and unknown parameters.
Song and Nelson (2013)	Bootstrap and variance decomposition	Independent univariate parametric distributions with known parametric families and unknown parameters.
Barton et al. (2014)	Meta-model-assisted bootstrapping (stochastic kriging meta-model)	Independent univariate parametric distributions with known parametric families and unknown parameters, and meta-model uncertainty can be ignored.
Xie et al. (2014a)	Bayesian credible interval assisted by meta-model (Gaussian process)	Independent univariate parametric distributions with known families and unknown parameters.
Xie et al. (2014b)	Meta-model-assisted bootstrapping (stochastic kriging meta-model) and Spearman's rank correlation	NORmal To Anything (NORTA) distribution [multivariate parametric distribution], which means unknown dependent inputs

	(unknown multivariate distribution), but known marginal distributions.

Table 3 shows a classification of the work discussed on Table 2 based on the assumptions made in them.

Table 3. Classification of work on simulation uncertainty quantification based on assumptions.

Type	**Description**	**Work**
Type A	Independent non-parametric distribution (empirical distribution)	Barton and Schruben (1993)
Type B	Independent parametric distributions (stationary univariate case)	Barton et al. (2010), Song and Nelson (2013), Barton et al. (2014), and Xie et al. (2014a)
Type C	Dependent parametric distributions (stationary multivariate case)	Xie et al. (2014b)
Type D	Time-varying independent parametric distributions (non-stationary univariate case)	None available to the best of our knowledge
Type E	Time-varying dependent parametric distributions (non-stationary multivariate case)	None available to the best of our knowledge

Under topic three, Shannon's entropy and mutual information are applied in conjunction with other methods such as contingency analysis, Tukey-Kramer multiple comparison test, and regression analysis for identifying the input parameters and understanding the experiment settings that contribute the most to the uncertainty in

simulation models. The results are compared to the results of the standard error of the mean and analysis of variance (ANOVA).

A general contribution of this dissertation is that although information theory has been widely recognized for its importance in the area of uncertainty and information quantification and feature classification, the theory has not been extensively applied in the simulation field yet. This dissertation investigates entropy measures as potential measures of uncertainty quantification in the simulation field and presents different potential applications of the measures for input parameter selection and for experiment planning.

Uncertainty quantification in simulation models is a well-studied topic through the use of meta-models, calibration techniques, and variance reduction techniques. This dissertation contributes to the body of knowledge of uncertainty quantification in the following manners: (i) by providing an empirical example to inform simulation stakeholders about the implications of uncertainty in the results of simulation results; (ii) by providing a method of uncertainty quantification in simulation models as a complement to the existing methods in the literature and which can be easily understood by the majority of the simulation stakeholders; and (iii) by being, to the best of our knowledge, the first work to provide applications of entropy measures for uncertainty quantification and experiment planning in the context of simulation models.

2. AN INVESTIGATION OF INFORMATION THEORY AS A METHOD FOR UNCERTAINTY QUANTIFICATION IN SIMULATION MODELS USING THE HISTOGRAM-BASED METHOD WITH STATIONARY UNIVARIATE DISTRIBUTIONS

2.1. Introduction

Decision-making is a part of everyday life. Examples include decisions about what, when, and from which supplier to buy a product, when to travel, which route to take, how much food to buy, and so on. These decisions made by both organizations and individuals usually involve different responses of interest that may or may not be related to each other. Although each of these decisions is likely different, there is a common concern among all of them: how uncertain is the available information and, consequently, what are the risks related to the decisions made?

When the system from which a decision needs to be made is composed of simple relationships, it may be possible to use mathematical methods to support the decision-making. However, when the relationships are too complex, which is the case of most real-world systems, the decision-making must be supported by approaches, such as simulation models (Law, 2007). Simulation is useful in supporting decision-making in different ways, such as: (i) by allowing the investigation of potential benefits brought by changes that are too costly to make in the real-world systems; (ii) by allowing the investigation of potential benefits and costs brought by proposed new systems; (iii) by allowing the investigation of

potential risks of scenarios that can be too dangerous for the system or society; (iv) by allowing the investigation of system bottlenecks; and so on.

Simulation model is, thus, a tool commonly used to support decision-making of complex systems. The ultimate goal of simulation models is to reduce the uncertainty in the decision-making process by providing more and reliable information about those complex systems. Despite being used to minimize uncertainties in the decision-making process, simulation models are driven by stochastic inputs and, hence, simulation-based estimates contain input estimation uncertainty as well as simulation estimation uncertainty (Xie et al., 2014a).

There are two main sources of uncertainty in simulation models: the input-uncertainty, due to fitting input distributions based on finite samples of real-world data; and the simulation-estimation error (or the simulation-sampling error), due to a finite amount of simulation effort (Xie et al., 2014a). Despite their uncertainties, simulation models can still effectively support decision-making and promote system improvements as long as simulation uncertainties are acknowledged (Christley et al., 2013).

In order to inform simulation stakeholders about the uncertainties present in simulation models, one needs to first quantify the uncertainty. The goal of uncertainty quantification is to identify and quantify the sources of error in simulation models and to assess their net and overall impact on the simulation results (DeVolder et al., 2002). Yet, an important question remains: how can one estimate uncertainty in simulation models? Within this context, one can see that the use and development of methods for uncertainty

quantification in simulation models are very important. The simulation uncertainty is commonly characterized by confidence intervals, and other methods have been developed in an attempt to quantify both the input and the simulation uncertainty. There is still more work to be done in this area. There are some unanswered questions, such as: (1) how uncertain are each one of the simulation inputs and outputs?, and (2) how much of the information in the simulation output reflects valid information from the input, and how much is noise?

In this dissertation, Shannon's entropy and mutual information are proposed as measures of simulation uncertainty. The main research question is: can entropy and mutual information measures quantify the uncertainty and, consequently, the information present in simulation models?

The analysis is restricted to simulation models using stationary univariate distributions. This restriction is justified based on the fact that: (i) stationary univariate distributions are the most common inputs in simulation models; and, (ii) this will give a good illustrative example for which closed-form solutions are available for assessment of the results.

The central contribution of this section is that it provides an analysis of Shannon's entropy and mutual information as measures of information and uncertainty in simulation models when using histogram-based method.

In Scheidegger, Banerjee, and Pereira (2018), the authors have proposed a framework for uncertainty quantification in simulation models, where they discussed the

sources and nature of uncertainty in simulation models, as well as the steps that should be followed to analyze and quantify uncertainty in simulation models. The authors also presented an application where mutual information was used as a measure of uncertainty in a mosquito-borne infectious disease simulation model using system dynamics. As the authors have highlighted in their conclusions, their work involved a simple application, which was a good starting point to show the potential of information theory for uncertainty quantification in simulation models. However, further investigation is needed to address some of the application's drawbacks and, as the authors have suggested, to explore future topics of investigation. Some of the drawbacks of Scheidegger et al. (2018) are: (i) the work used discrete empirical histogram estimate and Shannon's entropy for discrete variables, although their model was driven by continuous variables; (ii) the data was arbitrarily clustered into two bins with different sizes and the effect of number of bins and binwidth was not assessed; (iii) the potential of the measures as an uncertainty quantification method in simulation model was not discussed; and (iv) the authors only explored mutual information, but did not explore entropy measures.

This section is intended to be a continuation of Scheidegger et al. (2018), exploring some of the topics that were suggested for future work and addressing some of the drawbacks. There are six goals for this section of the dissertation: (1) discuss the challenges of computing entropy measures for continuous variables; (2) discuss the dependence of entropy on the binwidth; (3) investigate the entropy and mutual information as measure of uncertainty for different values of binwidth (fixed number of bins and

optimum number of bins); (4) investigate the measures for different normalization methods, different parameter values, and different contexts (different seeds for generating random numbers, constant work-in-progress (CONWIP), and addition of a third input -- travel time); (5) assess the potential of the measures as an uncertainty quantification method in simulation model; and, (6) compare the method when using histogram density estimate and discrete empirical histogram estimate. For a detailed list of the experiments performed and their configurations, please see Table 66 of the Appendix.

The rest of this section is organized as follows: section 2.2 provides a quick overview of studies in uncertainty quantification in simulation models and a discussion of entropy and mutual information, their challenges in the continuous case, and their dependence on the binwidth. Section 2.3 discusses the proposed application of entropy measures for uncertainty quantification in simulation models. Results and analyses are reported in section 2.4. Concluding remarks and future research directions are presented in section 2.5.

2.2. Background

2.2.1. Uncertainty quantification

Since the beginning of the 21st century, the topic of uncertainty quantification and propagation has been attracting the attention of simulation modelers from a wide variety of domains. Due to the increased importance of the topic, researchers now classify it as model verification, validation, and uncertainty quantification (VV&UQ), instead of the

previously well-known model verification and validation (V&V) paradigm (Roy & Oberkampf, 2011).

Every uncertainty analysis problem can be decomposed into input, model, and output uncertainty analysis (Marelli & Sudret, 2014). The input and output uncertainties are also known as extrinsic input-uncertainty and intrinsic output-uncertainty, respectively.

According to Song et al. (2014), the extrinsic input-uncertainty depends mainly on two factors: the amount of real-data available from which the input distribution parameters are estimated and the sensitivity of the response to those parameters. In other words, the input-uncertainty depends on (1) how accurately the input was modeled, and (2) how sensitive is the system response to the input model. Barton et al. (2010) mentioned that estimating the input-uncertainty may lead to a better balance between decision-making confidence and model results, and can also provide information about how much data must be collected to obtain model results at a desired level of accuracy.

The most common methods used in input-uncertainty propagation and quantification are: sampling-based methods, Bayesian methods, approximation methods, and meta-models or surrogate models (Barton, 2012; Baudin, Dutfoy, Iooss, & Popelin, 2015). Among the sampling-based methods, the most common ones are: minimum energy design, stratified sampling, direct resampling, bootstrap resampling, and meta-model assisted bootstrap (Barton, 2012). Approximation methods include first-order reliability method, second-order reliability method, and δ-methods (Marelli & Sudret, 2014).

Commonly used meta-models or surrogate models are: polynomial chaos expansion (also known as Wiener chaos expansion) and kriging (also known as Gaussian process regression). For a review of some of these methods, see Barton (2012). Graphical techniques and sensitivity analysis are also used for input-uncertainty propagation and quantification (Baudin et al., 2015; P. Chen, Quarteroni, & Rozza, 2013).

Regarding the intrinsic output-uncertainty, this uncertainty comes from the finite run length and the finite number of replications (Barton et al., 2010). The intrinsic output-uncertainty is already measured by all simulation software and it is characterized by confidence intervals on the performance measures (Barton et al., 2014; Song & Nelson, 2013). As in any sampling experiment, increasing the number of replications in a simulation project reduces the variance of the sample mean (Nelson, 1987a). However, increasing the number of replications may be too costly or not feasible due to time and computational resource constraints. Several techniques, such as antithetic variates (AV), control variates (CV), and common random numbers (CRN), have been developed to reduce the variance of simulation estimators without increasing the computational effort (Nelson, 1987b). Variance reduction techniques (VRTs) had their origins in Monte Carlo estimation and survey sampling around 1965 and 1975, respectively (Nelson, 1987a). Many simulation software offer built-in features that facilitate the execution of AV and CRN, but CV usually requires some additional software support (Nelson, 1987b).

There were not many methods found in the literature for model uncertainty analysis beyond model verification and validation.

Although there have been many methods applied to uncertainty quantification in a wide range of applications, there is no consolidated method so far (P. Chen et al., 2013). Besides, in general, the aforementioned methods focused on input-uncertainty quantification due to limited real-world data or interval quantification rather than total uncertainty quantification, and they require advanced mathematical and statistical knowledge that are not always possessed by every simulation stakeholder.

2.2.2. Entropy measures and mutual information

In 1948, Claude Shannon introduced the concept of entropy as a measure of information and uncertainty (Shannon, 1948). Shannon's theory accurately measures how much information can be transferred between different elements of a system and how uncertain is(are) the outcome(s) of the system (Stone, 2015). Shannon asked three main questions: (1) whether one could quantify the information produced by an information source or not; (2) the amount of choice involved in the message selection; and (3) how uncertain one would be about the outcome. In the context of uncertainty quantification in simulation models, Shannon's questions can be interpreted as: (1) how to quantify the information produced by simulation models?; (2) how to select the simulation model?; and, (3) how uncertain is(are) the simulation outcome(s)?

Related to Shannon's entropy measure, Stone (2015) asked another key question: how much of the entropy in the output reflects information in the input and how much is noise? This can be measured by the mutual information between X and Y, which is the average reduction in uncertainty about the value of X provided by the value of Y, and vice

versa. The mutual information (MI) measures the amount of information contained in a variable (or a group of variables) that helps to predict the system response. That is, MI is a symmetric measure that quantifies the statistical information shared between two variables (Ghosh, 2002). According to Dionisio, Menezes, and Mendes (2004), Kinney and Atwal (2014), and Haeri and Ebadzadeh (2014), MI is also a measure of linear and non-linear dependence among two variables. Despite the advantages of MI for quantifying information and the relationships between variables, its application, especially for continuous data, is not straightforward. MI requires an estimate of the probability distribution of the underlying data. How to compute this estimate in a way that does not bias the resulting MI remains an open problem (Kinney & Atwal, 2014).

In his work, Shannon discussed that if such a measure of information and uncertainty existed, it would require the following properties (Reza, 1961):

(i) Continuity: if the event probability is slightly changed, the associated measure of uncertainty or information should change accordingly in a continuous manner.

(ii) Symmetry: the measure must be functionally symmetric in relation to the probability set, i.e., the amount of information associated with a sequence of outcomes does not depend on the order in which those outcomes occur.

(iii) Extremal property: the maximum entropy is obtained when all the events are equiprobable.

(iv) Additivity: because entropy is nonnegative, partitioning the system into subevents cannot decrease the entropy of the system or the information associated with a set of outcomes is obtained by adding the information of individual outcomes.

Based on these properties, Shannon established that the only measure satisfying all the assumptions is represented by Equation 7, where k is a positive constant that only refers to the choice of the unit of measure and p_i is the probability of the i^{th} event.

$$H(X) = -k \sum_{i=1}^{n} p(x_i) \log p(x_i) \qquad \text{Equation 7}$$

In an analogous manner, but without any proof, Shannon defined in his paper that the entropy of a continuous variable with probability density function $f(x)$, known as differential entropy, is given by Equation 8.

$$H(X) = - \int_{x=-\infty}^{\infty} f(x) \log f(x)\, dx \qquad \text{Equation 8}$$

The above definition of differential entropy presents three main issues (Jaynes, 1968; Kittaneh, Khan, Akbar, & Bayoud, 2016): (i) it may be negative, while Shannon's entropy in the discrete case is always positive; (ii) it is not invariant under linear transformation; and, (iii) it is not a limit of Shannon's entropy of discrete approximations, which means that one is unlikely to estimate the differential entropy using the entropy of empirical distributions. Reza (1961) also added that the maximum entropy for a continuous variable does not occur when events are equiprobable anymore, as it was in the discrete case.

It is easy to show why the differential entropy may be negative. In the discrete case, the probability mass function is $0 \leq p(x_i) \leq 1$ and $\sum_{i=1}^{n} p(x_i) = 1$, which implies that $\log p(x_i) < 0$ or undefined when $p(x_i) = 0$ in which case $\log p(x_i)$ is considered to be equal to 0 in information theory. Because entropy is expressed as $-\sum_{i=1}^{m} p(x_i) \log p(x_i)$, entropy is always positive. On the other hand, in the continuous case the probability density function is $f(x) \geq 0 \; \forall \; x$ and $\int_{-\infty}^{\infty} f(x)\, dx = 1$. Because $f(x)$ can be greater than 1, $\log f(x)$ can be positive, in which case depending on the values of all x, the entropy may be negative.

Regarding the lack of invariance issue, Reza (1961) considered a new variable Y whose density function $\rho(y)$ was given in Equation 9.

$$\rho(y) = f(x)\left|\frac{dx}{dy}\right| \qquad \text{Equation 9}$$

The entropy associated with Y can be calculated using Equation 10.

$$\begin{aligned} H(Y) &= -\int_{-\infty}^{\infty} \rho(y) \log \rho(y) \\ &= -\int_{-\infty}^{\infty} \left[f(x)\left|\frac{dx}{dy}\right|\right] \log \left[f(x)\left|\frac{dx}{dy}\right|\right] dy \qquad \text{Equation 10} \\ &= H(X) + \int_{-\infty}^{\infty} f(x) \log \left|\frac{dx}{dy}\right| dx \end{aligned}$$

The linear transformation of the variable is represented by Equation 11.

$$Y = AX + B \qquad \text{Equation 11}$$

From Equation 10, it is possible to observe that the entropy of the continuous random variable Y will be changed by a constant $log|A|$ as shown in Equation 12.

$$H(Y) = H(X) + \log|A| \qquad \text{Equation 12}$$

According to Stone (2015) and Reza (1961), when the entropy of a continuous variable is discretized, that is, approximated by a discrete scheme, the entropy tends to infinity as the discretization is made finer and finer. Stone (2015) showed that by discretizing a continuous variable using a histogram in which each bin has a width equal to Δx and where the height of the histogram $p(x_i) = P_i/\Delta x$ was interpreted as a probability density. This discretization leads to Equation 13.

$$\begin{aligned} H(X^{\Delta}) &= -\sum_{i=1}^{m} P_i \log P_i = \sum_{i=1}^{m} p(x_i)\,\Delta x \log \frac{1}{p(x_i)\Delta x} \\ &= \sum_{i=1}^{m} p(x_i)\,\Delta x \left[\log \frac{1}{p(x_i)} + \log \frac{1}{\Delta x}\right] \\ &= \sum_{i=1}^{m} p(x_i)\,\Delta x \log \frac{1}{p(x_i)} + \sum_{i=1}^{m} P_i \log \frac{1}{\Delta x} \\ &= \sum_{i=1}^{m} p(x_i)\,\Delta x \log \frac{1}{p(x_i)} + \log \frac{1}{\Delta x} \end{aligned} \qquad \text{Equation 13}$$

Therefore, as the binwidth approaches zero, the first term on the right side of Equation 13 becomes an integral and the second term approaches infinity.

$$H(X) = \int_{x=-\infty}^{\infty} p(x) \log \frac{1}{p(x)}\, dx + \infty \qquad \text{Equation 14}$$

Stone (2015) argued that if all continuous variables had the same infinite entropy, this would not be useful. Although, in principle, each continuous variable can convey an

infinite amount of information, this is limited by the current capacity of the system and accuracy of measurement devices. In fact, the noise of measurement devices transforms continuous variables into discrete variables with m discriminable values, where m decreases as the noise increases (Stone, 2015). Therefore, researchers in the field of information theory agreed that the differential entropy of a continuous variable was the portion of Equation 14 that ignored the infinity and only included the "important" part, which agrees with the equation proposed by Shannon. However, important concerns remained in the field of information theory related to how to fix the issues presented in the differential entropy.

Methods and/or corrections have been proposed in the past to extend Shannon's entropy to continuous variables. Among the methods and corrections, three can be cited: (i) the cumulative residual entropy proposed in (Rao, Chen, Vemuri, & Wang, 2004); (ii) the estimation from a discrete approximation using histogram-based method and binwidth adjustment (Stone, 2015); and, (iii) passing to the limit from a discrete distribution and using an invariant measure (Jaynes, 1957, 1968).

In the correction mentioned in Stone (2015), when using the histogram-based method and equal binwidth, the differential entropy can be estimated using Equation 15. By using this method, entropy can be finitely estimated using a discrete scheme. However, the issues of the negativity and invariance still persist.

$$H_{dif}(X^{\Delta}) \approx \sum_i p(x_i)\,\Delta x \log\frac{1}{p(x_i)} = \sum_i P_i \log\frac{\Delta x}{P_i}$$

$$= \sum_i P_i \log\frac{1}{P_i} + \sum_i P_i \log \Delta x = \qquad \text{Equation 15}$$

$$= \left[\sum_i P_i \log\frac{1}{P_i}\right] - \log\frac{1}{\Delta x} = H(X^{\Delta}) - \log\frac{1}{\Delta x}$$

When the entropy is discretized, it is known that its maximum value is achieved when all signals (or datapoints) are equally likely. That is, $H_{Max}(X) = log_2\, n$. When the binwidth is calculated based on the number of bins, we have: $\Delta x = \frac{\max(x_i) - \min(x_i)}{n}$. When the data is normalized between 0 and 1, we have: $\max(x_i) = 1$, $\min(x_i) = 0$, and, consequently, $\Delta x = \frac{1}{n}$. Therefore, $H_{dif}(X^{\Delta}) \approx H(X^{\Delta}) - \log\frac{1}{\Delta x} \leq H_{Max}(X) - log_2\, n \leq 0$, when the binwidth is based on the number of bins.

Jaynes (1957) proposed as a correction to the continuous case to pass it to the limit from a discrete distribution by using Equation 16.

$$H(X) = -\int p(x) \log\left[\frac{p(x)}{m(x)}\right] dx \qquad \text{Equation 16}$$

where $m(x)$ is an invariant measure proportional to the limiting density of discrete points.

A question remains regarding what the measure $m(x)$ should be. As pointed out in Jaynes (1968), if the parameter space is not the result of any limiting process, the conclusions will depend on the measure chosen. The measure $m(x)$ also has an impact in

the maximum entropy principle, which in the discrete case occurred when all the events were equiprobable.

Reza (1961) mentioned that due to the difficulties in the continuous case, entropies have no direct interpretation with respect to the information or uncertainty. However, the author highlighted that the mutual information preserves its properties and, therefore, its relevance. The mutual information is nonnegative, invariant under linear transformation, and the problem of infinity disappears as the measure involves the difference between two entropies (Reza, 1961; Stone, 2015). More specifically, mutual information is bounded (Egnal & Daniilidis, 2000).

Reza (1961) and Egnal and Daniilidis (2000) mentioned that just as the definition of random variables could be extended from one-dimension to two-dimension, the definition of entropy could also be extended for joint and conditional entropy. The conditional entropy $H(Y|X)$ reflects the uncertainty of Y when X is known or the noise and is given by Equation 17 (Stone, 2015). In other words, the more Y depends on X, the lower the conditional entropy. The joint entropy $H(X,Y)$ reflects the average information of the system associated with the pair X and Y and is given by Equation 18 (Egnal & Daniilidis, 2000; Stone, 2015). The mutual information is given by Equation 19 (Stone, 2015).

$$H(Y|X) = -\sum_{i=1}^{m_x}\sum_{j=1}^{m_y} p(x_i, y_j)\, log\, p(y_j|x_i) \qquad \text{Equation 17}$$

$$H(X,Y) = -\sum_{i=1}^{m_x}\sum_{j=1}^{m_y} p(x_i, y_j)\log p(x_i, y_j) = H(X) + H(Y|X) \quad \text{Equation 18}$$

$$I(X;Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y|X) \quad \text{Equation 19}$$

As the entropy of variables can vary significantly, a normalized mutual information (NMI) is desirable for easier interpretation and comparisons (Estévez, Tesmer, Perez, & Zurada, 2009; Strehl & Ghosh, 2002). Another reason for normalizing is that MI is biased towards multi-binned variables, which means that MI increases with the increase in the number of bins used to calculate the entropy of the variables (Estévez et al., 2009; Strehl & Ghosh, 2002). Different normalization methods have been suggested in the literature: (1) $NMI_{arith} = 2\frac{I(X;Y)}{H(X)+H(Y)}$; (2) $NMI_{geom} = \frac{I(X;Y)}{\sqrt{H(X)\times H(Y)}}$; (3) $NMI_{joint} = IQR = \frac{I(X;Y)}{\mathrm{H(X,Y)}}$; and, (4) $NMI_{theor} = NMI_{log} = \frac{I(X;Y)}{\max(\max(\mathrm{H(X)}),\max(\mathrm{H(Y)}))} = \frac{I(X;Y)}{\max(\log n_X, \log n_Y)}$. These methods are discussed in Ghosh (2002), Hill, Batchelor, Holden, and Hawkes (2001), Principe, Xu, Zhao, and Fisher (2000), and Strehl and Ghosh (2002).

Another difficulty during the calculation of entropy measures in the continuous case is that the evaluation of the integral in Equation 8 requires numerical integration and it is computationally inefficient (Beirlant, Dudewicz, Györfi, & Van der Meulen, 1997). To overcome this difficulty, the following approximations are given in Xiong, Faes, and Ivanov (2017) and Steuer, Kurths, Daub, Weise, and Selbig (2002): (1) $\widehat{H}(X) =$

$-\frac{1}{n}\sum_{i=1}^{n}\log\hat{f}(X_i)$; (2) $\widehat{H}(X,Y) = -\frac{1}{n}\sum_{i=1}^{n}\log\hat{f}(X_i,Y_i)$; (3)$\widehat{H}(X|Y) = -\frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{\hat{f}(X_i,Y_i)}{\hat{f}(Y_i)}\right]$; and, (4) $\hat{I}(X;Y) = \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{\hat{f}(X_i,Y_i)}{\hat{f}(X_i)\hat{f}(Y_i)}\right]$.

2.2.3. Histogram-based method and binwidth selection

Despite the advantages of information theory for quantifying information and the relationships between variables, its application is not simple, especially for continuous data. Entropy measures and MI require an estimate of the probability distribution of the underlying data. How to compute this estimate in a way that does not bias the resulting measures remains an open problem (Kinney & Atwal, 2014). There are three main non-parametric approaches discussed in the literature: the histogram-based method, the kernel-based method, and the k-nearest neighbors (KNN) distance method (Xiong et al., 2017). Legg, Rosin, Marshall, and Morgan (2013) mentioned that the histogram-based method is the most commonly used method for density estimation.

In the histogram-based method, the probability functions are approximated by means of histograms where the data is divided into bins and the number of elements in each bin is counted (Dionisio et al., 2004).

For equally spaced bins, the empirical distribution function is given by Equation 20 (Castro, 2015; Koshkin, 2014; Waterman & Whiteman, 1978).

$$\hat{p}_j(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{I}\{x_i \in [t_j, t_{j+1})\}\ for\ x_j \in B_j, j = 1, \dots, k \qquad \text{Equation 20}$$

where $B_j = [t_j, t_{j+1})$ denotes the j^{th} bin of a total of k bins, and $\mathbf{I}(.)$ is the indicator function which is 1 if $x_i \in [t_j, t_{j+1})$, and 0 otherwise.

$\hat{p}_j(x)$ converges in probability to $p_j(x)$ as $n \to \infty$ and its mean estimator is unbiased; however, $\hat{p}_j(x)$ is not continuous but a staircase function, which is a disadvantage if one wants to use it to approximate a continuous random variable (Castro, 2015; Koshkin, 2014). For the statistic $\hat{p}_j(x)$ to be meaningful when the random variable is continuous, a probability density estimator must be used (Foutz, 1980).

In the histogram density estimation, the distribution is estimated by counting the number of data points that are in each bin and assigning to that bin a probability equal to the number of points it contains divided by the total number of data points and the binwidth. When equally spaced bins are used, the histogram density can be estimated using Equation 21 (Pace, 1995).

$$\hat{f}^{hist}{}_j(x) = \frac{1}{nh}\sum_{i=1}^{n} \mathbf{I}\{x_i \in [t_j, t_{j+1})\} \; for \; x_j \in B_j, j = 1, \dots, k \qquad \text{Equation 21}$$

where h is the binwidth, $B_j = [t_j, t_{j+1})$ denotes the j^{th} bin of a total of k bins, and $\mathbf{I}(.)$ is the indicator function which is 1 if $x_i \in [t_j, t_{j+1})$, and 0 otherwise.

Equation 21 can be extended to the multivariate case; however, Silverman (1986) recommended that as in any other multivariate procedure, one should normalize the data to avoid extreme differences of spread in various directions.

A drawback of the histogram-based method is that it is not a continuous function and it is not differentiable at the boundaries of the bins, which is undesirable to estimate a

continuous probability density function. On the other hand, the histogram-based method is very simple, easy to understand, computationally very efficient, and it makes few assumptions about the probability function it is trying to estimate (Egnal & Daniilidis, 2000). MI estimates based on this method are often called naïve estimates as they may overestimate or underestimate $I(X;\ Y)$, the mutual information between X and Y (Dionisio et al., 2004; Kinney & Atwal, 2014). Moreover, selecting the bin size (or length) is the main source of error. Kinney and Atwal (2014) argued that this is not a problem in the large data limit, because the probabilities can be determined to arbitrary accuracy as $n \rightarrow \infty$.

As seen from Equation 21, the histogram method is dependent on the choice of the binwidth and also of the choice of the origin of the bin (or the start point of the bin) (Härdle, Müller, Sperlich, & Werwatz, 2012; Kanazawa, 1993; Xiong et al., 2017). As a result, a common question when constructing a histogram is the size of the binwidth. Selecting the bin length (or binwidth) is the main source of error on entropy measures using the histogram-based method. Legg et al. (2013) stated that the binwidth choice is also critical in the effectiveness of the mutual information.

According to Egnal and Daniilidis (2000), Wand (1997), and Kanazawa (1993), selecting the binwidth parameter is the most important choice in the histogram method because it controls the trade-off between bias and variance or oversmoothing and undersmoothing (Scott, 1979). If the number of bins is too large, the variance is high, but the bias is low. On the other hand, if the number of bins is low, the variance is low, but

the bias is high. Similarly, a small binwidth results in a rough, undersmoothed, histogram and a large binwidth results in a single block, oversmoothed histogram (Scott, 1979; Wand, 1997). Ideally, the binwidth should balance the variance and bias, and should be chosen so that the histogram displays the essential structure of the data, without giving too much weight to the dataset at hand (Egnal & Daniilidis, 2000; Scott, 1979; Wand, 1997).

Wand (1997) mentioned that there are not many methods for choosing the starting point of the bins apart from looking at different shifted histograms with the same binwidth, but there are a number of proposed methods to theoretically determine the appropriate binwidth of equal length. Some methods take a look at the data at hand and based on the data decide the binwidth, while others try to predict the binwidth based on prior knowledge and assumptions, and others select a measure of discrepancy between the histogram and the density and, then, asymptotically minimize the expected value of the measure (Egnal & Daniilidis, 2000; Kanazawa, 1993). Legg et al. (2013) and Birgé and Rozenholc (2006) stated that although there are many methods, there is not a consensus on how to choose the binwidth because none of the methods have been completely proved to be better than the others. While rules of thumb are very simple and do not aim at any optimality property, more sophisticated rules are based on asymptotic estimates and, in general, do not warrant good performance for small sample size (Birgé & Rozenholc, 2006).

For practical efficiency, the majority of the methods for binwidth selection is based on dividing the range of sampled data into k equally sized bins. Optimizing the number of bins among regular partitions, or equally sized bins, is computationally easy and fast.

While the use of irregular partitions (or differently sized bins) reduce the bias, optimizing for irregular partitions increases the complexity of the selection problem significantly (Birgé & Rozenholc, 2006).

A simple and probably the oldest rule of thumb for the number of bins was proposed by Sturges in 1926 and is given by Equation 22 (Legg et al., 2013; Scott, 2015; Sturges, 1926). The rule is based on the properties of the data and assumes the data is normal (Legg et al., 2013). The binwidth is calculated by dividing the data range by the total number of bins (Egnal & Daniilidis, 2000). Although simple, Sturges' rule is shown to lead to oversmoothed histogram, especially for large samples (Legg et al., 2013; Wand, 1997).

$$k = 1 + log_2 n \qquad \text{Equation 22}$$

where n is the sample size and k is the number of bins.

In 1976, Doane proposed a variation of Sturges' rule to allow for skewness, which is also known to lead to oversmoothed histograms (Hyndman, 1995; Wand, 1997). Another famous rule was proposed by Scott in 1979 (Scott, 1979). Scott used the mean integrated squared error (MISE) to obtain the asymptotically optimal binwidth for normal data (Kanazawa, 1993). Scott's rule is given by Equation 23 (Scott, 2015; Wand, 1997).

$$h = 3.49\sigma n^{-1/3} \qquad \text{Equation 23}$$

where σ is the sample standard deviation. The rule proposed by Scott led to better large sample performance of the histogram, but is not consistent itself (Wand, 1997). Scott

also proposed a modification of his rule for varying degrees of skewness and kurtosis (Scott, 1979, 2015).

Freedman and Diaconis' (FD) rule given by Equation 24 uses the interquartile range (IQR) of the data instead of the standard deviation as Scott's rule (Hyndman, 1995; Legg et al., 2013).

$$h = 2\,IQR\,n^{-1/3} \qquad \text{Equation 24}$$

Other rules include: Devroye and Györfi's rule, Taylor's rule, and Hall's rule (Kanazawa, 1993; Legg et al., 2013). Although there are many rules for binwidth selection, based on a literature review performed by Legg et al. (2013), there is little discussion regarding joint histogram binwidth selection and the impact of the binwidth on the MI has not been fully investigated in the literature so far.

2.3. Material and methods

One of the main goals of a simulation model is to support and improve decision-making. In order to avoid backfiring, the results of a simulation model must be as accurate as possible and the uncertainty of the results must be acknowledged. In this context, there are two questions to be answered regarding simulation models: how can one quantify the uncertainty of the results and, consequently, the information produced by simulation models?; and, how uncertain is(are) the simulation outcome(s)? The use of information-based measures, namely Shannon's entropy and mutual information, is proposed to quantify uncertainty and information in simulation models. The main goal is to determine

whether these entropy measures are adequate for the purpose of quantifying uncertainty and information in simulation models.

For simplification, the analysis is focused on measuring the total uncertainty of each output and determining the extent of this uncertainty that can be attributed to each of the simulation generated inputs. In other words, how much of this uncertainty can be attributed to the inputs that are actually generated by the simulation software; e.g., the pseudo-random numbers. The reason for this choice is that, if simulation modelers have to choose, they should be more worried about the uncertainty of the simulation model responses than any other uncertainty and how the inputs impact this uncertainty. According to Song et al. (2014), the input uncertainty depends not only on the amount of real-data available from which the input distribution parameters are estimated but also on the sensitivity of the response to those parameters. In an analogous way, it is possible to say that the output uncertainty depends not only on the finite amount of simulation effort, but also on the sensitivity of the simulation response to the inputs, which justifies investigating how much of the output uncertainty can be attributed to the simulation generated inputs. Therefore, identifying the inputs that have a greater impact on the outputs can provide information about where one should spend more effort in reducing the uncertainty in order to reduce the total output uncertainty and to obtain the desired level of simulation output accuracy.

Entropy measures are used to quantify the aforementioned idea. Figure 5 shows a schematic of the proposed use of entropy measures to quantify uncertainty in simulation

models. As can be seen in Figure 5, the simulation outputs are impacted by different sources of noise (or uncertainty), namely: system noise, input modeling noise, and simulation noise. The system noise is an inherent source of uncertainty and can never be eliminated. As such, system noise will also be part of the total uncertainty of the simulation output results as uncertainty is propagated throughout the system. The assumption is that the values of the real system inputs and outputs are unknown (shown as dashed arrow). This is the most common situation in the real world, especially with respect to the outputs. Therefore, entropy measures are applied to quantify the total uncertainty of the simulation outputs ($H(Y)$), the uncertainty of the simulation generated inputs ($H(X)$), and how much of the simulation output uncertainty can be attributed to each of the simulation generated inputs ($I(X;Y)$). According to Wijaya, Sarno, and Zulaika (2017), the greater the mutual information between two variables, the greater the impact these variables have on each other.

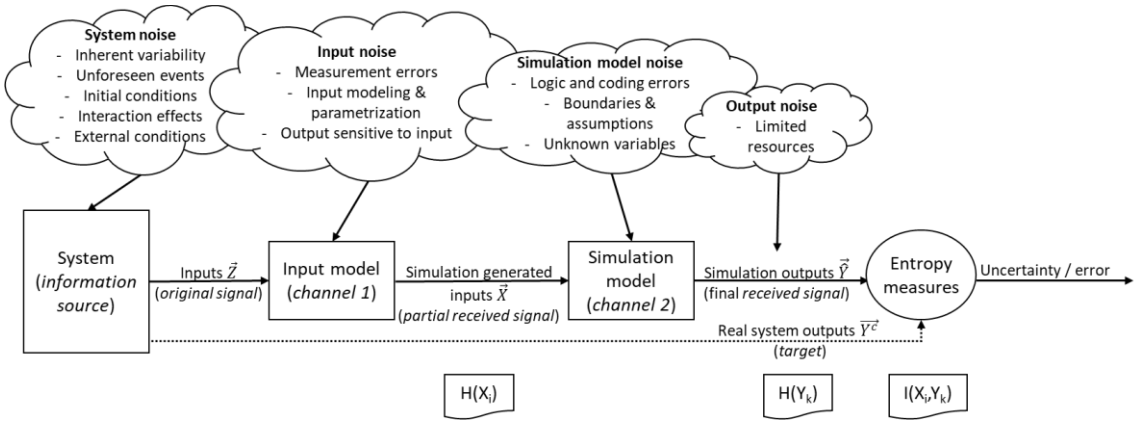

Figure 5. Schematic of the use of entropy measures to quantify uncertainty in simulation models.

There are many types of input models that can be used in simulation models, such as: univariate input distributions, time-dependent (non-stationary) inputs, and multivariate input distributions. Here, the focus is on investigating the application of entropy measures in simulation models that consist of l independent stationary univariate input processes and the average estimate of p outputs of interest. Specifically, the adequacy of the proposed method is investigated through an $M/M/s$ illustrative example with two simulation generated input processes and two outputs of interest, namely: $\vec{X} = [X_1: arrival\ process,\ X_2 - service\ process]$ and $\vec{\hat{Y}} = [\hat{Y}_1: average\ number\ of\ customers\ in\ the\ system,$ $\hat{Y}_2: average\ time\ in\ the\ system]$. Experiments on $M/G/1$ and $G/G/s$ systems are run to investigate how the entropy measures behave in different systems. However, only $M/M/s$ and $M/G/1$ systems are used to investigate the potential of the measures as an uncertainty quantification method due to the availability of their exact closed-form solutions.

The experiments are run using Simio® University Enterprise Edition v 10.165. The experiments include different traffic intensities and several number of replications. The model consists of a single source of arrivals, single queue, and s servers providing the same service. After being served, entities leave the system. Balking and reneging were

not considered in the model. The notation follows the $A/S/s$ Kendall's notation, where: A represents the arrival process, S the service time, and s the number of servers. M is used for memoryless distributions.

There are a few challenges when applying entropy measures for continuous variables. Taking these challenges and important points into consideration, entropy and mutual information measures are investigated as potential and adequate measures for uncertainty quantification in simulation models by: (i) investigating the entropy and MI as measures of uncertainty for different values of binwidth; (ii) investigating the mutual information measure for different normalization methods; (iii) investigating the entropy and MI measures for different parameter values and different systems; (iv) assessing the potential of the measures as an uncertainty quantification method for $M/M/s$ and $M/G/1$ systems for which exact closed-form solutions are available; and, (v) comparing the method when using histogram density estimate and discrete empirical histogram estimate. Each one of these goals are discussed in detail below.

The entropy measures are calculated using the histogram-based method, where the probability density functions are approximated by means of histograms by dividing the continuous data into bins and by counting the number of elements in each bin as given by Equation 21. The histogram-based method was chosen due to its computational efficiency and for being commonly used and wider acceptance in academia. After computing the probability density function, entropy and mutual information measures are then calculated using approximations, instead of using numerical integration. However, applying entropy

measures to continuous variables brings a difficulty of interpretability due to the challenges previously discussed. In this case, a solution is needed to overcome the issue. A few solution approaches have been proposed in the literature. Some issues still arise when applying these solutions to continuous variables and, as a result, a new solution is proposed.

In order to apply Equation 21 to estimate the probability density function using the histogram-method, one needs to define the number of bins to be used and the binwidth. Here, the binwidth is determined in two different ways: (1) fixed number of bins; and (2) optimum number of bins, which uses formulas proposed in the literature to determine the binwidth. For fixed number of bins, the following number of bins are investigated: 2; 5; 10; 25; 50; 100; 200; 500; 1,000; and, 2,000. These values were chosen in order to investigate a large range for number of bins, which, in turn, leads to large and small binwidths. The binwidth is calculated by taking the data range and dividing it by the number of bins, as given by Equation 25. The start position of the first bin is given by the minimum of the data range (Equation 26) and, based on that, all the remaining bins are determined, as given by Equation 27. Following this calculation procedure, the end of the last bin must coincide with the maximum of the data range (Equation 28). For optimum number of bins, three different formulas are investigated: Sturges' rule, Scott's rule, and FD's rule.

$$h = \frac{\max\left(\mathrm{x}_i\right) - min\ \left(\mathrm{x}_i\right)}{k};\ i = 1, \dots, n \qquad \text{Equation 25}$$

$$t_1 = \min(\mathrm{x}_i);\ i = 1, \dots, n \qquad \text{Equation 26}$$

$$t_j = t_{j-1} + h;\ j = 2, \dots, k \qquad \text{Equation 27}$$

$$t_{k+1} = \max(\mathrm{x}_i);\ i = 1, \dots, n; \qquad \text{Equation 28}$$

where h is the binwidth, n is the sample size, k is the number of bins, t_1 is the start of the 1^{st} bin, t_j is the start of the j^{th} bin, and t_{k+1} is the end of the k^{th} bin.

Four normalization methods of mutual information are investigated: (1) NMI_{arith}, (2) NMI_{geom}, (3) NMI_{joint}, and, (4) NMI_{theor}.

$H(X)$ is the average uncertainty of variable X, and $I(X;Y)$, the mutual information between X and Y, is the average reduction in uncertainty of the value of Y provided by the value of X or the amount of information shared between these variables, their dependence. Considering these definitions, different approaches are used to assess the potential of the measures as a method of uncertainty quantification in a simulation model. The potential of the entropy are assessed by comparing the entropy results with the sum of absolute error (SAE), sum of squares error (SSE), mean absolute error (MAE), and mean squared error (MSE) of the inputs and the outputs of the system. SSE and SAE can be calculated given that the inputs were defined by the user and here they are assumed to be correct, and the theoretical outputs can be calculated using results from queueing system theory. SSE and SAE were chosen because these measures of error and uncertainty are widely accepted in the literature. The potential of the mutual information are assessed by comparing the mutual information results with three different measures of dependence between variables:

distance correlation (Székely & Rizzo, 2009; Székely, Rizzo, & Bakirov, 2007), Pearson correlation, and $R^2{}_{adj}$ (or adjusted coefficient of determination). Distance correlation was chosen because it is a measure of dependence between two random variables that is able to capture both linear and non-linear association (Székely et al., 2007). Pearson correlation and $R^2{}_{adj}$ were chosen for being widely accepted in the literature. The downside of these latter measures is that they only capture the linear association between the random variables.

Finally, the last analysis performed is a comparison of the results when using histogram density estimate, as given by Equation 21 and discrete empirical function, as given by Equation 20, which is a gross approximation of the simulation inputs and outputs as it considers the inputs and outputs as discrete variables.

2.4. Results and discussion

2.4.1. Challenges encountered while applying entropy measures for continuous variables and method proposed to overcome the issues

The solutions proposed in the literature were used to overcome the challenges faced while applying entropy measures to continuous variables. The first solution chosen was the one provided in Equation 15, where the differential entropy is approximated by calculating the discretized entropy and by adjusting with a correction dependent on the binwidth. According to Stone (2015), this approximation eliminates the problem of infinity. However, this method does not eliminate the issue of negative values, which leads to a difficulty in interpretability and for performing comparison.

It is known that for the discrete case, Shannon's entropy is maximum when all the n events are equiprobable, which is given by Equation 29 .

$$Max[H(X^{\Delta})] = -\log_2 p = -\log_2 \frac{1}{n} = \log_2 n \qquad \text{Equation 29}$$

Equation 15 can be rewritten as Equation 30 when binwidth is calculated based on a fixed number of bins.

$$\begin{aligned} H_{dif}(X^{\Delta}) &= H(X^{\Delta}) + \log_2 \frac{\max(x_i) - \min(x_i)}{n} \\ &= H(X^{\Delta}) + \log_2[\max(x_i) - \min(x_i)] - \log_2 n \end{aligned} \qquad \text{Equation 30}$$

From Equation 29, one can see that the maximum value $H(X^{\Delta})$ is $\log_2 n$. Therefore, from Equation 30 it is easy to see that depending on the value of the data range and depending on the actual value of the entropy, which will be at most $\log_2 n$, the entropy may be negative and the difficulty in interpretability remains. As recommended by Silverman (1986), when the data is normalized between 0 and 1 before calculating the entropy, we have $H_{dif}(X^{\Delta}) = H(X^{\Delta}) + \log_2[1] - \log_2 n \leq Max[H(X^{\Delta})] - \log_2 n \leq 0$.

Next, the solution proposed by Jaynes (1957) and shown in Equation 16 was investigated. In this method, it is important to define the invariant function $m(x)$. While Jaynes has not provided any suggestion for the function, there are a few suggestions made in the literature: $m(x) = sup\left[\hat{f}(x)\right]$ and $m(x) = E\left[\hat{f}(x)\right]$ (Awad & Alawneh, 1987; Kittaneh et al., 2016).

One can immediately see that $m(x) = E[\hat{f}(x)]$ is not a good choice as $\hat{f}(x)$ may be greater than $E[\hat{f}(x)]$. In this case, $\frac{\hat{f}(x)}{m(x)} > 1 \Rightarrow log_2 \frac{\hat{f}(x)}{m(x)} > 0$ and, hence, entropy is negative.

The suggestion to use $m(x) = sup[\hat{f}(x)]$ does not have this issue. For $m(x) = sup[\hat{f}(x)]$, $\hat{f}(x)$ is never greater than $m(x)$. In this case, $0 \leq \frac{\hat{f}(x)}{m(x)} \leq 1 \Rightarrow log_2 \frac{\hat{f}(x)}{m(x)} \leq 0$ and, hence, entropy is always positive. However, there remain some issues of interpretability when compared to the discrete case as shown in Table 4 below.

Table 4. Issues of interpretability of entropy in the continuous case when using $\boldsymbol{m(x) = sup[f(x)]}$.

Scenario	**Discrete case result (Shannon's definition)**	**Continuous case result (using $\boldsymbol{m(x) = sup[f(x)]}$)**	**Interpretability issue?**
Events are all equiprobable (or all probability density function values are equal)	Maximum entropy	$\hat{f}(x) = \mathrm{m(x)} \Rightarrow \frac{\hat{f}(x)}{m(x)} = 1$ $\Rightarrow log_2 \frac{\hat{f}(x)}{m(x)} = 0$ $\Rightarrow\ entropy = 0$	Yes
Events are either equiprobable or surely will not occur/impossible (probability density function values are equal or are 0)	Maximum entropy	$\hat{f}(x) = \mathrm{m(x)} \Rightarrow \frac{\hat{f}(x)}{m(x)} = 1$ $\Rightarrow log_2 \frac{\hat{f}(x)}{m(x)} = 0$ $\Rightarrow\ entropy = 0$	Yes
All events will surely not occur/impossible	Entropy is equal to 0	Entropy is equal to 0	No

(probability density function values are 0)			
Event will surely occur (certainty)	Entropy is equal to 0	How to represent certainty in the nonparametric continuous case? If $\hat{f}_i(x) = \sup \hat{f}(x)$ and $\hat{f}_j(x) = 0; j \neq i$, should this be considered the certainty case? If yes: $\hat{f}(x) = \mathrm{m(x)} \Rightarrow \frac{\hat{f}(x)}{m(x)} = 1$ $\Rightarrow log_2 \frac{\hat{f}(x)}{m(x)} = 0$ $\Rightarrow entropy = 0$ However, one could interpret that certainty does not exist in the continuous case as one can never be sure about an event occurring in the continuous case ($P(X = x) = 0$).	Maybe

As shown in Table 4, when using Jaynes' method and $m(x) = sup\left[\hat{f}(x)\right]$, there are still difficulties in interpreting the entropy results from the discrete case to the continuous case. This may not be an issue depending on the application. However, here, the goal is to quantify uncertainty in simulation models and to be able to identify the inputs that contribute the most to the uncertainty and the simulation response that users should be most careful while making decisions due to its uncertainty. To achieve this goal, it is

important to be able to compare entropy and MI measures among themselves, which is difficult to perform when entropy measures do not have a lower bound and can be negative and/or do not have an upper bound and can be infinite.

Observing this challenge and recalling what Silverman (1986) suggested to do when one is working with multivariate estimates, the following solution was proposed to work with entropy in the continuous case. Silverman (1986) recommended that data should be normalized to avoid extreme differences of spread in the variables. So, the solution here involves normalizing the data in a way that does not only avoid the differences of spread, but also guarantees that $0 \leq \hat{f}_i(x) < 1, \forall i$.

The binwidth for fixed number of bins is calculated using Equation 31.

$$binwidth_b = h_b = \frac{\max(x_i) - \min(x_i)}{k_b}; b = 1, \dots, m \qquad \text{Equation 31}$$

$$\hat{f}^{hist}{}_j(x)_b = \frac{1}{nh_b} \sum_{i=1}^{n} \mathbf{I}\{x_i \in [t_j, t_{j+1})\} = \frac{1}{nh_b}\ ;\ j = 1, \dots, k_b \qquad \text{Equation 32}$$

$$0 \leq \frac{1}{nh_b} < 1 \Rightarrow 0 \leq \frac{k_b}{n[\max(x_i) - \min(x_i)]} < 1, b = 1, \dots, m \qquad \text{Equation 33}$$

$$\hat{f}^{hist}{}_j(x)_b = \frac{n}{nh_b} = \frac{1}{h_b};\ \ j = 1, \dots, k_b \qquad \text{Equation 34}$$

$$0 \leq \frac{1}{h_b} < 1 \Rightarrow 0 \leq \frac{k_b}{\max(x_i) - \min(x_i)} < 1;\ b = 1, \dots, m \qquad \text{Equation 35}$$

where b is the number of different number of bins that are investigated, h_b is the binwidth of the b^{th} number of bin being investigated, k_b is the total number of bins of the b^{th} number of bin being investigated, and n is the number of data points.

Equation 31 is derived from Equation 25. Equation 21 shows how to estimate the probability density function based on the histogram-method. One extreme case is to have only one data point in the bin and the other extreme is to have all the n data points in the bin, which from Equation 21 leads to Equation 32 and Equation 34, respectively. Replacing Equation 31 in both Equation 32 and Equation 34 leads to Equation 33 and Equation 35. Because Equation 35 is more restrictive, as long as Equation 35 is satisfied, Equation 33 is also satisfied ($\max(x_i) - \min(x_i) > 0$, $n > 0$, and $k_b > 0$).

Therefore, the data normalization is determined by finding the data range that satisfies Equation 35 for all b different number of bins simultaneously. Obviously, there will be multiple solutions, but any solution can be picked. Now, all the probability density function values are between 0 and less than 1. This enables calculation of entropy measures in a similar way as the discrete case and similar interpretation can be made. It is worth noting that different from the discrete case the interval excludes 1, because whenever the probability mass function is equal to 1 all the other values are 0 and the entropy is 0. However, in the continuous case, the probability density function may be equal to 1 with other values different than 0. In this case, the entropy should not be 0, as there is some uncertainty in the system regarding the events. To avoid this issue of the extreme case equal to 1 ($log_2 1 = 0$), 1 was excluded from the interval.

By using the aforementioned alternative, the issues discussed in Table 4 are now minimized to the scenario where the event will surely occur and to the fact that the entropy

will not be maximum when the density functions are equal anymore, but it will at least not be null, as shown in Table 5.

Table 5. Issues of interpretability of entropy in the continuous case when using proposed data normalization method.

Scenario	**Discrete case result (Shannon's definition)**	**Continuous case result (using proposed normalization alternative)**	**Interpretability issue?**
Events are all equiprobable (or all probability density function values are equal)	Maximum entropy	$0 < \widehat{f}(x) < 1 \Rightarrow entropy \neq 0$	Not the same result from discrete variable, but entropy is not 0 which allows for comparison with entropy results of other variables or experiments.
Events are either equiprobable or surely will not occur/impossible (probability density function values are equal or are 0)	Maximum entropy	$0 < \hat{f}(x) < 1 \Rightarrow entropy \neq 0$ Entropy will only be equal to 0 if all the probability density function values are equal to 0, which in this case the result will match the discrete variable result.	Not the same result from discrete variable, but entropy is not 0 which allows for comparison with entropy results of other variables or experiments.
All events will surely not occur/impossible	Entropy is equal to 0	Entropy is equal to 0	No

(probability density function values are 0)			
Event will surely occur (certainty)	Entropy is equal to 0	How to represent certainty in the nonparametric continuous case? If $\hat{f}_i(x) = \sup \hat{f}(x)$ and $\hat{f}_j(x) = 0; j \neq i$, should this be considered the certainty case? If yes, there is an issue, as this case will not lead to entropy equal to 0. In fact, there will never be entropy equal to 0, except when all probability density function values are equal to 0. However, one could think that certainty does not exist in the continuous case as one can never be sure about an event occurring in the continuous case ($P(X = x) = 0$).	Maybe

As can be seen, the procedure described above can be implemented when using fixed number of bins, but, in general, it cannot be implemented when using formulas that calculate the optimum number of bins or binwidth, because the binwidth is not a function of the data range. This is an advantage of the fixed number of bins over the latter.

Lemma 1. When using fixed number of bins, changing the data normalization does not change the placement of data into bins because the binwidth is recalculated accordingly.

It is possible to calculate the binwidth and start- and end-points of the bins as given by Equation 36 to Equation 39.

$$h_{original\ data} = \frac{\max_{original\ data} - min_{original\ data}}{k} \quad \text{Equation 36}$$

$$t_{1_{original\ data}} = min_{original\ data}; \quad \text{Equation 37}$$

$$t_{j_{original\ data}} = t_{j-1} + h;\ j = 2, \dots, k \quad \text{Equation 38}$$

$$t_{k+1_{original\ data}} = \max_{original\ data} \quad \text{Equation 39}$$

where h is the binwidth, k is the number of bins, t_j is the start of the j^{th} bin, t_{k+1} is the end of the k^{th} bin, and $B_j = [t_j, t_{j+1})$ denotes the j^{th} bin of a total of k bins.

The data normalization is performed using the approach shown in Equation 40.

$$x_{new} = \frac{x_{original\ data} - min_{original\ data}}{max_{original\ data} - min_{original\ data}} \times (max_{desired} - min_{desired}) + min_{desired} \quad \text{Equation 40}$$

Now, suppose there is a data point x_1 that is contained in the first bin, as given by Equation 41.

$$x_{1_{original\ data}} \in B_1 = [t_1, t_2)$$

$$= \left[min_{original\ data}, min_{original\ data} + h\right)$$

$$= \left[min_{original\ data}, min_{original\ data} + \frac{max_{original\ data} - min_{original\ data}}{k}\right) \quad \text{Equation 41}$$

$$= \left[min_{original\ data}, \frac{max_{original\ data} + (k-1)\, min_{original\ data}}{k}\right)$$

$x_{1_{original\ data}}$ can be normalized using Equation 40. Let us now consider the extremes. If $x_{1_{original\ data}} = min_{original\ data}$, we have:

$$x_{new} = \frac{min_{original\ data} - min_{original\ data}}{max_{original\ data} - min_{original\ data}} \times (max_{desired} - min_{desired}) + min_{desired} \quad \text{Equation 42}$$

$$= min_{desired}$$

Now, if the other extreme is considered $x_{1_{original\ data}} = \frac{max_{original\ data} + (k-1)\, min_{original\ data}}{k}$, we have:

$$x_{new}$$

$$= \frac{\frac{max_{original\ data} + (k-1)\ min_{original\ data}}{k} - min_{original\ data}}{max_{original\ data} - min_{original\ data}}$$

$$\times (max_{desired} - min_{desired}) + min_{desired}$$

Equation 43

$$= \frac{\frac{max_{original\ data} - min_{original\ data}}{k}}{max_{original\ data} - min_{original\ data}}$$

$$\times (max_{desired} - min_{desired}) + min_{desired}$$

$$= \frac{max_{desired} - min_{desired}}{k} + min_{desired}$$

Therefore, it is possible to see that $x_{new} \in B_{1_{new}} = [min_{desired}, \frac{max_{desired} - min_{desired}}{k} + min_{desired}) = [min_{desired}, \mathrm{h} + min_{desired}) = [t_{1_{new}}, t_{2_{new}})$ and the normalization did not change the placement of data into bins because the binwidth was recalculated accordingly using the data range.

2.4.2. The impact of different binwidths and different normalization methods on entropy and mutual information measures

The discussion in this and the following sections are based on the results of the experiments, which were detailed in section 2.3 and listed in Table 66 of the Appendix. For simplification, the experiments are referred by their numbers.

It is worth mentioning again that entropy and MI measures are calculated in three different ways: (1) using the histogram-based method with fixed number of bins and probability density function based on the data normalization approach; (2) using the

histogram-based method with optimum number of bins and probability density function; and, (3) using the histogram-based method with fixed number of bins and discrete empirical distribution function (discrete assumption). The third case was performed for comparison as it was the method adopted in Scheidegger et al. (2018).

The analysis of Figure 6 shows that in the histogram-based method with fixed number of bins and probability density function based on data normalization, the entropy and MI measures tend to decrease with the increase in the number of bins (or decrease in the binwidth) for the same number of replications. This is an important observation, as this contradicts what is mentioned in the literature. The exception is when the number of bins is small, between 2 and 10. In this range, there is no clear pattern and sometimes the entropy and MI measures increase with the increase in the number of bins. The reason for the observation here to be different from what has been found in the literature is that, although entropy and MI measures are very useful in many fields, in the majority of applications the measures have been applied for discrete variables only or considering that continuous variables can be approximated using discrete probability functions and, hence, binwidth is not taken into account in the histogram estimate. This will be discussed in more detail later in this section.

From Figure 6, one can also see that the entropy and MI measures tend to increase with the increase in the number of replications for the same number of bins (or binwidth). The amplitude of the increase is higher for number of bins that are not too large nor too small. When the number of bins is very large or very small, the amplitude of the increase

becomes smaller. The exception is when the number of bins is equal to 2. In this case, the entropy and MI measures do not increase or decrease considerably with the increase in the number of the replications.

The number of bins used to calculate the entropy measures can be considered as the level of accuracy one wants to obtain or one is interested in (i.e., number of bins = 1/targeted accuracy). This idea is similar to the indifference zone concept, where two response values separated by less than the specified indifference zone value are considered to be statistically equivalent. In a low number of bins scenario, the number of bins data can be interpreted as: "the stakeholder is not concerned about a high level of accuracy". Therefore, running the first few replications should bring a great amount of information about the inputs and outputs, and consequently, a great reduction in the uncertainty of the output provided by the input ($I(X;Y)$), because initially, the stakeholder had no information about the system. The next set of replications should still bring valuable information about the system, if one does not have all the information. However, because the stakeholder is not interested in a high level of accuracy, all the information about the system should be gathered faster than when compared to a larger number of bins, which means that the curve should stabilize faster (smaller slope of the curve) and at some point running more replications would not yield any significant reduction in the uncertainty than what was already provided during the initial number of replications. On the other extreme, when the stakeholder is interested in very high level of accuracy, the initial replications may not yield enough information or reduce the uncertainty, but more replications would

result in a greater reduction of the uncertainty or more information would be available until stability is reached. This stability will be reached at a later point than for lower number of bins (lower accuracy). This could be seen as oversmoothing and undersmoothing, respectively. Of course, in between the extremes, there are mid-range values of the number of bins where the initial number of replications would provide great reduction in uncertainty. This corresponds to a steep slope at first, and then as more replications are run less information is obtained, and running more replications would not be economically efficient anymore. The fact that the entropy and MI decrease with the increase in the number of bins is expected from a measure of information and uncertainty. When there is an emphasis on a greater level of detail or accuracy, the same number of replications (or the same amount of the data) should be able to provide less information or less reduction in uncertainty.

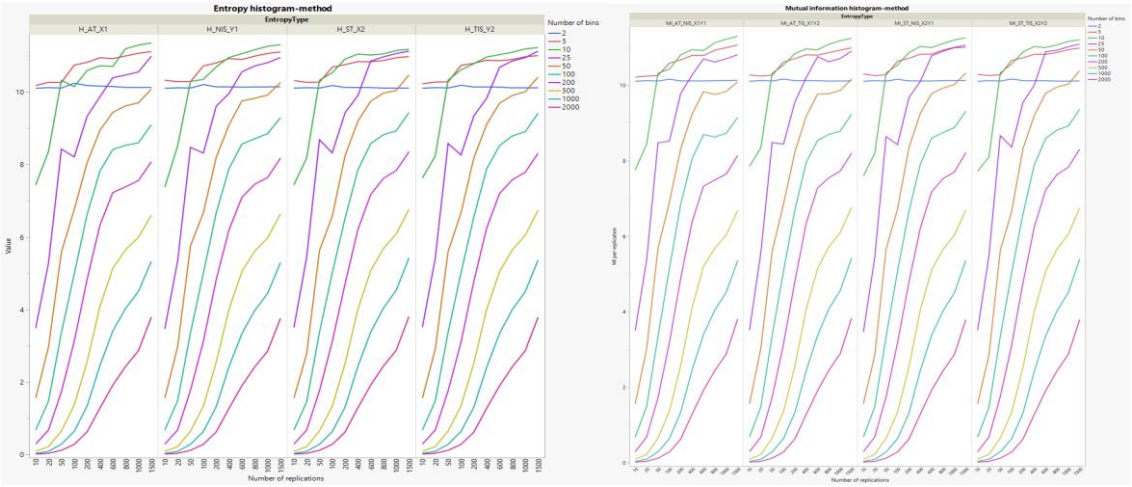

Figure 6. Average of entropy and MI measures per number of bins using histogram-based method with fixed number of bins and probability density function (experiments #1 to #350).

There are three rules that are considered for obtaining the optimum number of bins: Sturges', Scott's, and FD's rules. These rules take the amount of data available and/or the data characteristics, such as data dispersion, to calculate the optimum number of bins or binwidth. For the experiments in this study, this resulted in low number of bins (or larger binwidth) in all the cases. Figure 7 shows the impact of the binwidth on the entropy and MI measures based on the optimum number of bins rules. The conclusions are similar to the ones obtained from Figure 6 for low number of bins (i.e., between 5 and 10): there is a tendency for the entropy and MI measures to increase with the increase in the number of replications; however, this tendency is not as prominent as in Figure 6 anymore because the number of bins parameter in Figure 7 is not fixed as it was in Figure 6.

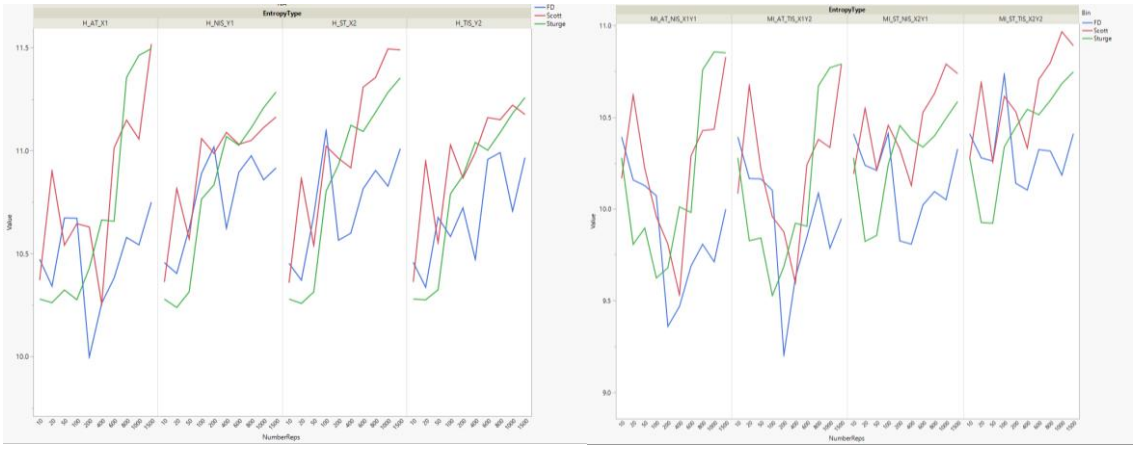

Figure 7. Average of entropy and MI measures per number of bins using histogram-based method with optimum number of bins and probability density function (experiments #1 to #350).

Figure 8 shows the entropy and MI measures per number of bins when the histogram-based method is used with discrete empirical distribution function. From Figure 8, one can see that as mentioned in the literature, the entropy and MI measures tend to increase with the increase in the number of bins (or decrease in the binwidth) for the same number of replications. This is different than what was observed for the histogram-based method with probability density function. In the literature it is acknowledged that the number of bins have an impact in the entropy and MI measures and, consequently, should be taken into account. The difference between the histogram-based method with probability density function and the histogram-based method with discrete empirical distribution appears to be due to the binwidth being taken into account in the histogram estimate of the first method, but not in the latter. This goes in agreement with what Stone (2015) has stated in his book about the estimated entropy of a discretized continuous variable increasing with the decrease of the binwidth. With respect to the number of replications, the entropy and MI measures present a similar behavior to what was observed in the histogram-based method with probability density function. However, here, for low number of bins, the entropy and MI measures tend to be nearly constant or decrease with the increase in the number of replications, while when using probability density function the entropy and MI tend to be nearly constant or increase. For larger number of bins, the

entropy and MI measures first increase, when the number of replications is not too large. When the number of replications become too large, the entropy and MI measures stop increasing and stabilize. Different than the probability density function, the amplitude of the increase is greater for larger number of bins.

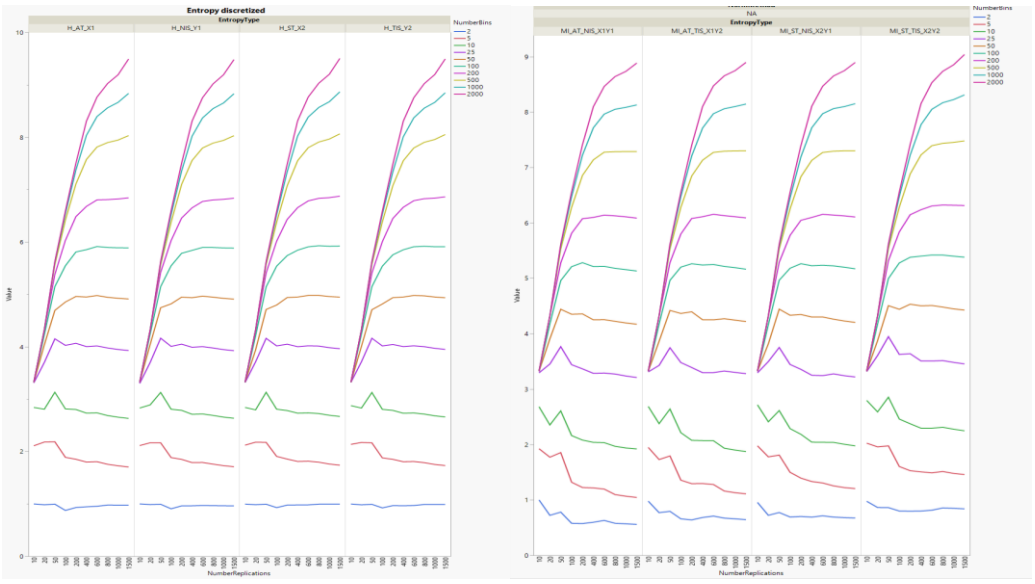

Figure 8. Average of entropy and MI measures per number of bins using histogram-based method with fixed number of bins and discrete empirical distribution (experiments #1 to #350).

As suggested in the literature, entropy and MI measures should be normalized to eliminate the impact of the bins on the measures. The entropy was normalized by its maximum: $NH(X) = \frac{H(X)}{\max(H(X))}$ or $NH(Y) = \frac{H(Y)}{\max(H(Y))}$. The MI was normalized using the four different normalization formulas discussed in section 2.3. Figure 9 shows the results of the normalized entropy and MI measures per number of bins when the histogram-based

method is used with discrete empirical distribution function. From Figure 9, one can see that after normalization the entropy and NMI_{theor} measures present a behavior similar to the observations in Figure 6. That is, the measures tend to decrease with the increase in the number of bins, with exceptions when the number of bins is between 2 and 25 and a clear pattern is not observed. The measures also tend to increase with the increase in the number of replications. The exception again occurs when the number of bins is smaller and between 2 and 25. This exception did not occur in the results shown in Figure 6. Based on the other three normalization methods ($NMI_{arith}, NMI_{geom}, NMI_{joint}$), the MI measures still tend to increase with the increase in the number of bins as they did before normalization. However, after normalization using these three methods (arith, geom, and joint) the MI measures tend to decrease with the increase in the number of replications.

There are a few important points to highlight. First, the normalized entropy and the NMI_{theor} measures present a behavior similar to the entropy and the MI measures calculated using the histogram-based method with probability density function and fixed bins. Second, the difference in the behavior between the NMI_{theor} and the $NMI_{arith}, NMI_{geom}\ and\ NMI_{joint}$ can be explained by the normalization method adopted. NMI_{theor} was normalized using the theoretical maximum value of the entropy measures, which is fixed regardless of the number of replications and changes with number of bins. $NMI_{arith}, NMI_{geom}\ and\ NMI_{joint}$ were normalized using the real maximum value (from the data) of the entropy measures, which varies based on the number of replications and number of bins. The entropy measures increase with the

increase in the number of replications and number of bins as observed in the results shown in Figure 8.

It is also important noting that the results of the MI normalization methods appear very similar but are not identical, as shown in Figure 9. This is due to the fact that the same number of bins were used to calculate the entropy of all the inputs and outputs, and therefore the maximum value of each of the inputs and the outputs is the same. The recommendation is to use the same number of bins for all inputs and outputs whenever using entropy measures as a method for uncertainty quantification in simulation models. The idea comes from blocking in the design of experiments. By using the same number of bins, the effect of the bins on the entropy values and, hence, on the uncertainty value is being minimized.

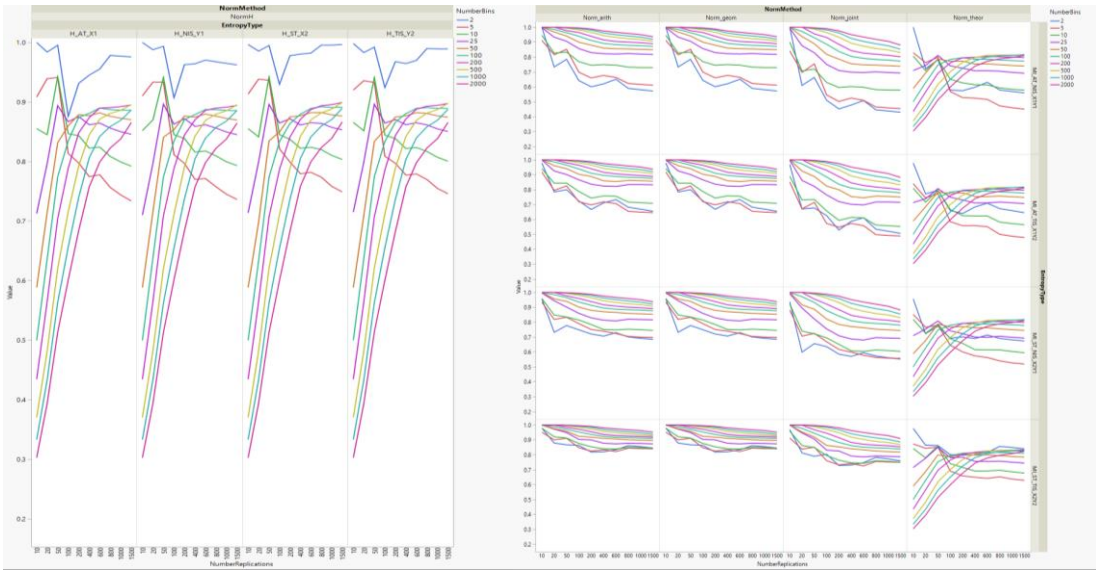

Figure 9. Average of normalized entropy and MI measures per number of bins using histogram-based method with fixed number of bins and discrete empirical distribution (experiments #1 to #350).

Figure 10 shows the entropy and MI measures using the histogram-based method with probability density function per queue model. Figure 10 is similar to Figure 6, but it allows to see that the values of the entropy and MI measures are not identical among themselves and vary based on the input and output, as well as queue model and other factors. However, when the average over all the experiments is considered as in Figure 6, the entropy and MI measures appear to be almost identical.

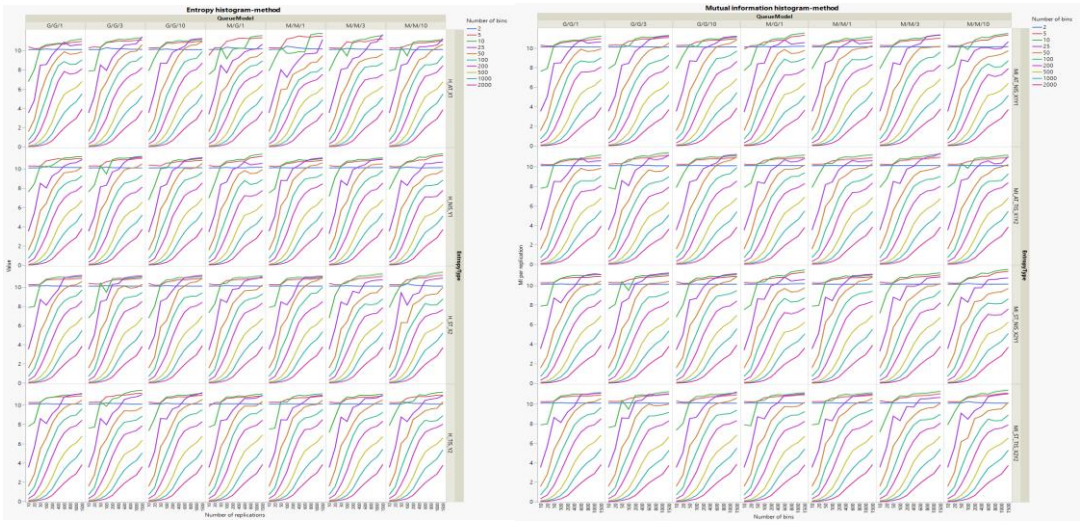

Figure 10. Average of entropy and MI measures per number of bins per queue model using histogram-based method with fixed number of bins and probability density function (experiments #1 to #350).

2.4.3. The impact of different traffic intensities, different seeds, different parameter values, and different systems on entropy and mutual information measures

It is important to evaluate the appropriateness of entropy and MI as measures of uncertainty quantification in simulation models. The role of different traffic intensities, different seeds, different parameter values, and different systems impacting the measures are investigated here.

In the queue example used in this study, it is known that the uncertainty of the input X_1 must be equal among the different traffic intensities because the same input model and fixed seed were used in the simulation model. Based on the results obtained for the histogram-based method using probability density function, which can be seen in Figure 11, the entropy of X_1 was equal among the different traffic intensities which indicates that the entropy measure is possibly accurately measuring the information or uncertainty of X_1.

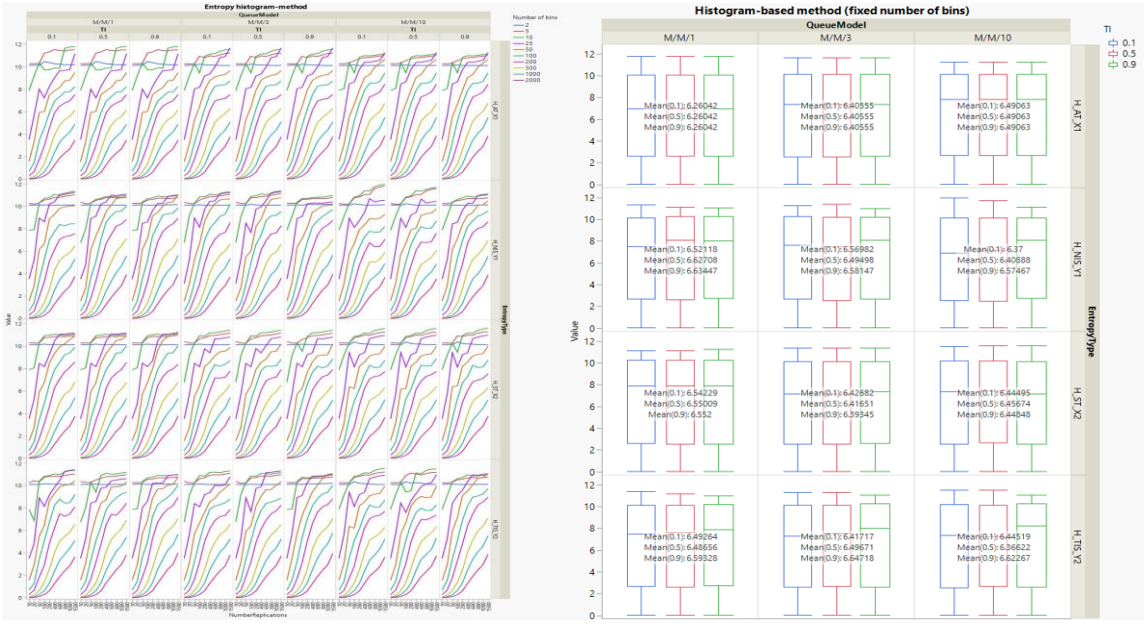

Figure 11. Entropy measures per queue model per traffic-intensity using histogram-based method with fixed number of bins and probability density function (experiments #1 to #90).

Although a fixed seed was also used for input X_2, it is not appropriate to expect that X_2 should also have equal entropy among different traffic intensities. The impacts of X_2 depend on how the system was modeled. In the approach adopted in this dissertation, changes in traffic intensities were modeled by changing the capacity of the only existing server, instead of adding or eliminating servers. By doing so, even though the seed of the service time input, X_2, is fixed, the generated inputs X_2 changed and, thus, its entropy should not remain the same among the different traffic intensities. Reviewing Figure 11, it is possible to observe that the entropy of X_2 was able to capture some of the differences among the different traffic intensities, which also indicates that the entropy measure is correctly measuring the information or uncertainty of the simulation generated input X_2.

Similar results as the ones obtained for the histogram-based method with probability density function and fixed number of bins were obtained for histogram-based method with probability density function and optimum number of bins, and also for the histogram-based method with discrete empirical distribution function and fixed number of bins. This is shown in Figure 12.

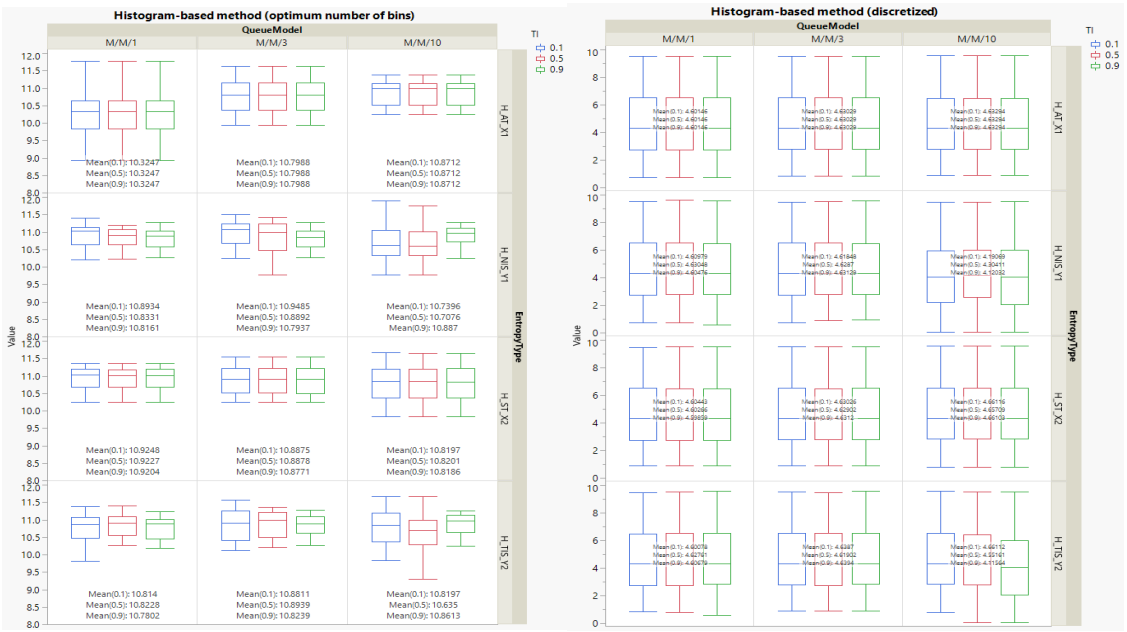

Figure 12. Entropy measures per queue model per traffic-intensity (experiments #1 to #90) using histogram-based method with: (a) optimum number of bins with probability density function (left-side) and (b) fixed number of bins with discrete empirical distribution (right-side).

Another point is that one could expect the reduction in uncertainty in the output provided by the input to be different in a high traffic intensity system than in a low traffic intensity system. That is, the reduction in uncertainty of the average time in system ($\hat{Y}_2$)

provided by the input service time (X_2), or the mutual information $I(X_2;Y_2)$, is expected to be different among the different traffic intensities. In systems that are headed towards an unstable state, it is natural to expect the service time to have a different impact on the average time in the system than in more stable systems. As shown in Figure 13, this was indeed observed in the MI measures. Similar results were obtained when using optimum number of bins or discrete empirical distribution.

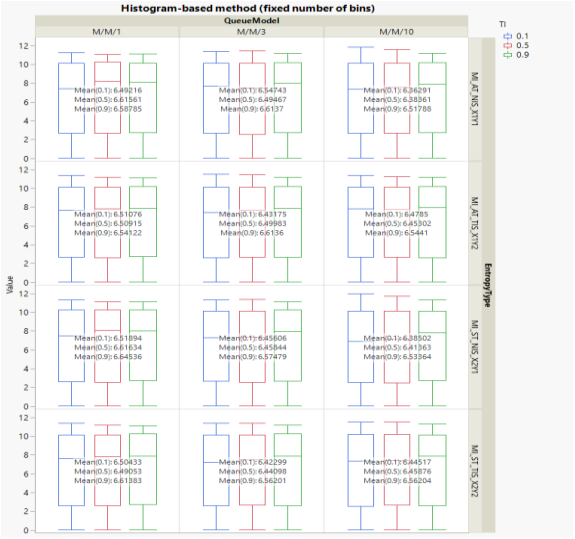

Figure 13. MI measures per queue model per traffic-intensity using histogram-based method with fixed number of bins and probability density function (experiments #1 to #90).

As mentioned earlier, experiments with different seeds were run to investigate the appropriateness of entropy and MI as measures of uncertainty quantification. As shown in Table 66, experiments #351 to #460 and #461 to #470 correspond to the initial

experiments numbered #1 to #110 and #271 to #280, respectively, but with a different group of seeds (here, named group "seed 2", the specific seed numbers used are shown in Table 66). That is, the same experiment configuration was kept: same interarrival time, same service time, same number of servers, but different groups of seeds were used for the parameters. Similarly, experiments #471 to #580 and #581 to #590 correspond to the initial experiments numbered #1 to #110 and #271 to #280, respectively, but with another seed (here, named group "seed 3").

As shown in Figure 14 and Figure 15, the entropy measures and mutual information vary based on the group of seeds used. Although one may initially not expect this to occur because the seeds are fixed, a good measure of uncertainty should indeed vary based on the seeds being used. Entropy is a measure of the information or uncertainty of the inputs and outputs. For different seeds, there are different uncertainties. Although these differences should not be large because the same input model (or distribution) is being used, the values cannot be identical either; otherwise, it would mean that exactly the same information was observed, which is unlikely when using different seeds for generating pseudo-random numbers. Therefore, by using a different group of seeds, one should expect different entropy values for the inputs, different entropy values for the outputs, and, consequently, different MI values as the results in Figure 14 and Figure 15 show. From Figure 14, one can also see that regardless of the group of seeds, the entropy of X_1 is equal among the different traffic intensities. Although it is not shown here, similar results were obtained for the entropy and MI measures calculated using the histogram-

based method with optimum number of bins and the histogram-based method with fixed number of bins and discrete empirical distribution.

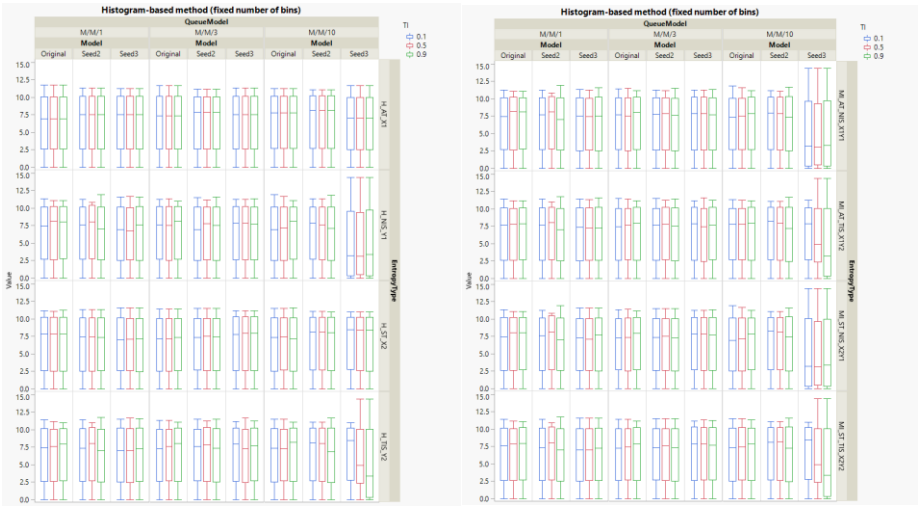

Figure 14. Entropy and MI measures per queue model per traffic-intensity per seed using histogram-based method with fixed number of bins and probability density function.

From Figure 16 it is possible to make another interesting observation about the impact of the group of seeds on the entropy and MI measures. As shown in Figure 16, "seed 3" had a different impact in the entropy of $\hat{Y}_1$ and $\hat{Y}_2$ when compared to both the original group of seeds and "seed 2", and "seed 3" had a similar impact in the entropy of X_1 and X_2 when compared to both the original group of seeds and "seed 2". When analyzing the MI, the impact of "seed 3" occurred in every MI measure, however the impact appeared to be greater in $I(X_1;Y_1)$ and $I(X_2;Y_1)$ than in $I(X_1;Y_2)$ and $I(X_2;Y_2)$, when compared to the original group of seeds and "seed 2". This is also observed in the

results shown in Figure 17. From Figure 17, one can also see that entropy and MI measures may be used to investigate the adequacy of seeds when running simulation models. While one expects to observe some differences in the uncertainty values for using different groups of seeds (different data), one does not expect a large difference as the input models are the same and in the long-run the data should be similar. Therefore, if the difference is large, as it is for group "seed 3" and especially for larger traffic intensity, this may possibly indicate an issue with the group of seeds.

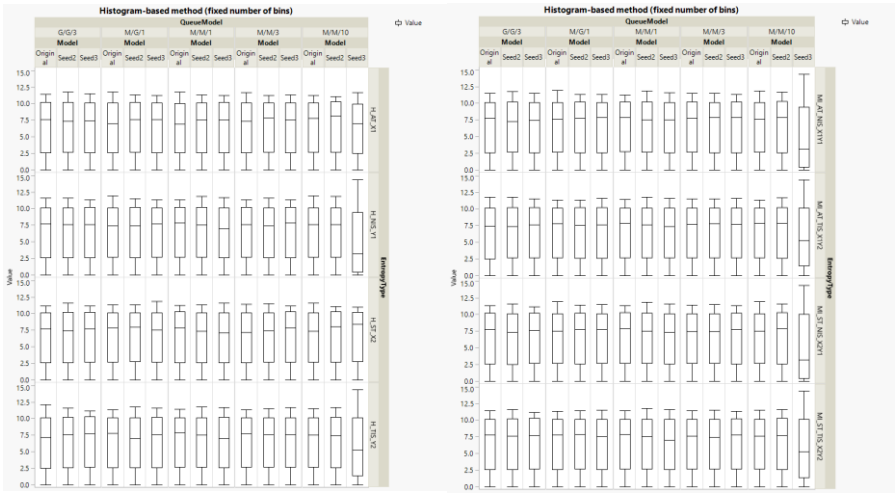

Figure 15. Entropy and MI measures per queue model per seed using histogram-based method with fixed number of bins and probability density function.

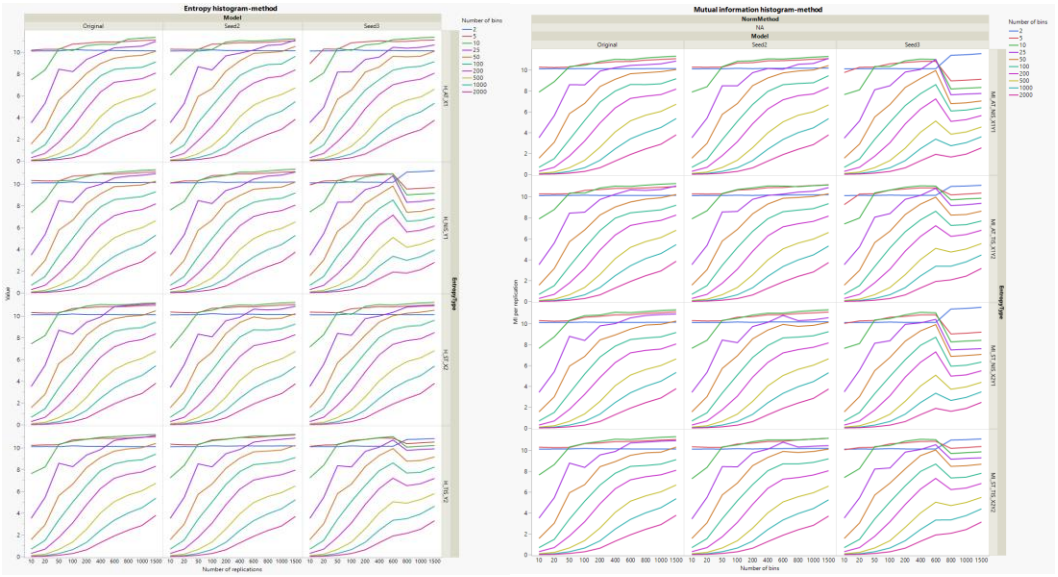

Figure 16. Entropy and MI measures per replication per number of bins using histogram-based method with fixed number of bins and probability density function.

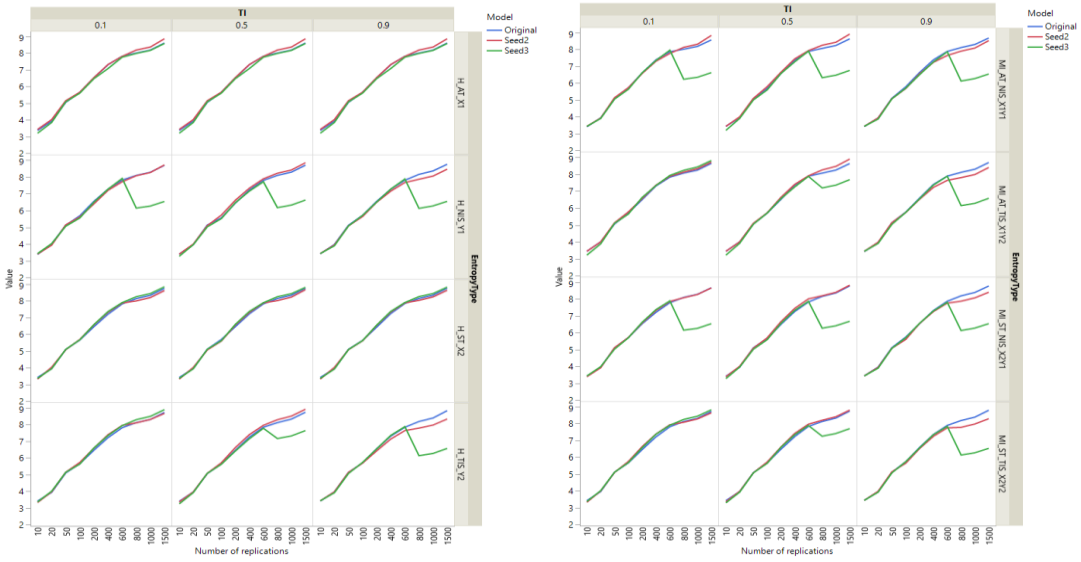

Figure 17. Entropy and MI measures per traffic-intensity per model using histogram-based method with fixed number of bins and probability density function.

To investigate the appropriateness of entropy and MI as measures of uncertainty quantification, experiments with different parameter values were also run. Experiments #591 to #700 and #701 to #710 correspond to the initial experiments numbered #1 to #110 and #271 to #280, respectively, but with different parameter values (here, named "number 2", the specific parameter values used in the experiments are given in Table 66). Experiments #711 to #820 and #821 to #830 correspond to the initial experiments numbered #1 to #110 and #271 to #280, respectively, but with different parameter values (here, named "number 3").

From Figure 18 one can see that even though different values of inputs X_1 and X_2 were used, the entropy of X_1 was different among the different experiments: "original", "number 2", and "number 3", but it was still equal among the different traffic intensities within each group of experiments as expected for a fixed seed. Also, from Figure 18, one can see that the traffic intensity appears to not have a clear relation to the uncertainty of the outputs, as the uncertainty either increases or decreases based on the queue model and that changes in the system configurations led to different values of uncertainty. Similar results were obtained for the entropy and MI measures calculated using the histogram-based method with optimum number of bins and the histogram-based method with fixed number of bins and discrete empirical distribution, as shown in Figure 19 and Figure 20.

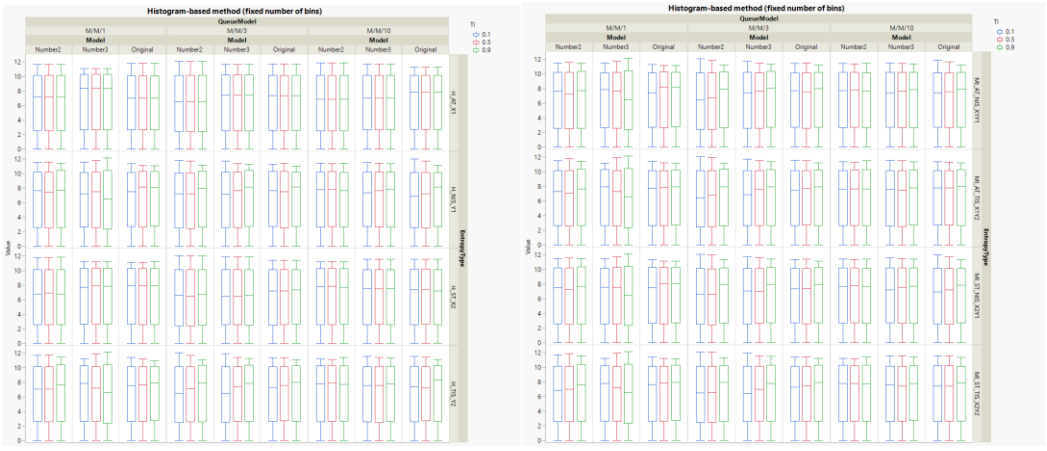

Figure 18. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using histogram-based method with fixed number of bins and probability density function.

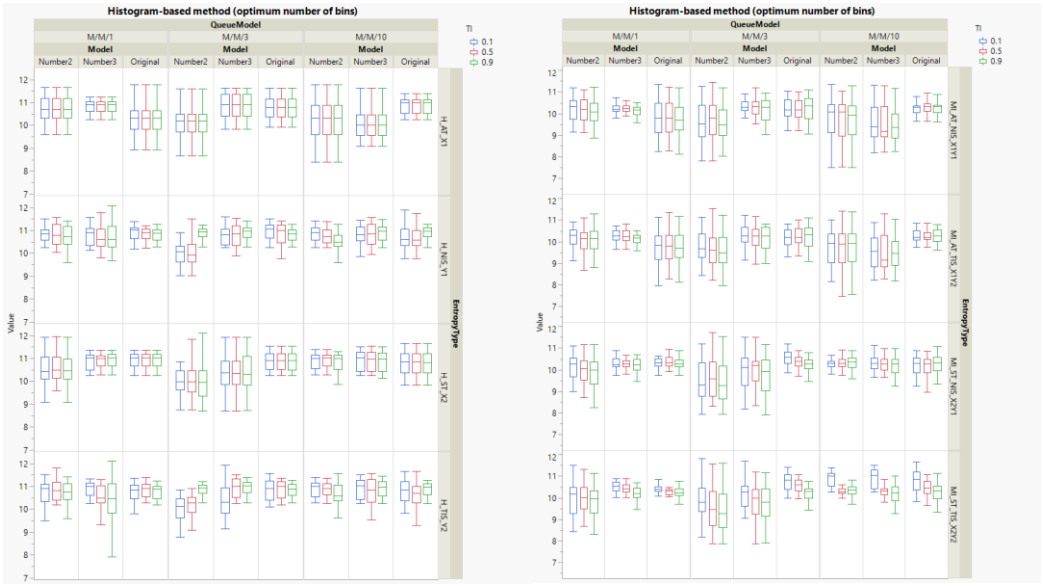

Figure 19. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using histogram-based method with optimum number of bins and probability density function.

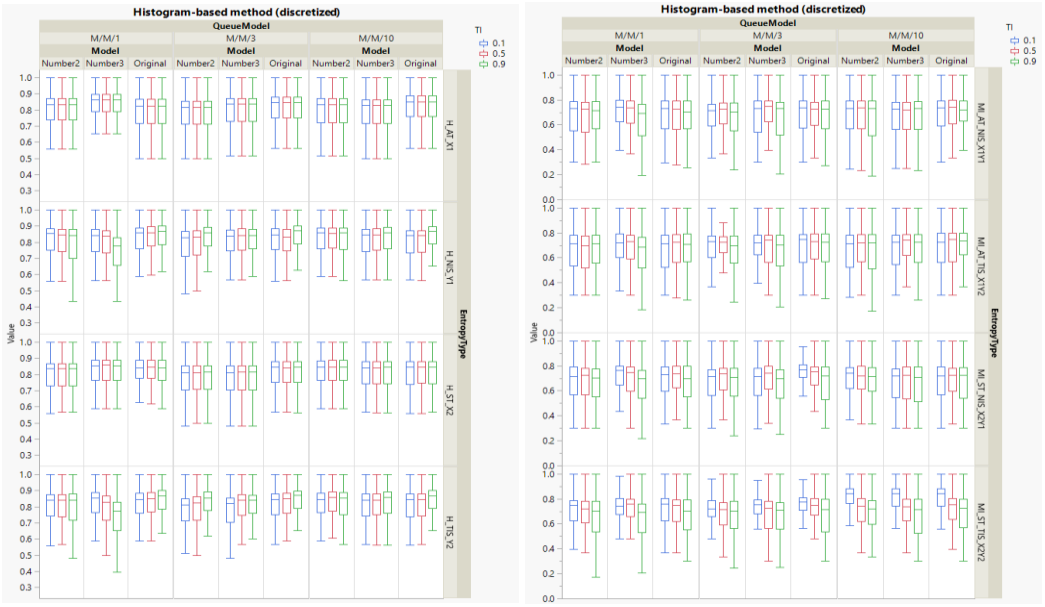

Figure 20. Normalized entropy and $\boldsymbol{NMI_{theor}}$ measures per queue model per traffic-intensity per parameter value using histogram-based method with fixed number of bins and discrete empirical distribution.

Finally, a last group of experiments was run to investigate whether the entropy and MI measures would be able to capture the uncertainty of different systems. It was decided to investigate two additional systems: a CONWIP system and the addition of a third input parameter, namely travel time, in the queue system.

A total of 100 CONWIP experiments were run. Experiments #831 to #930 in Table 66 correspond to the initial experiments numbered #1 to #100, with same service time but constant work in progress. CONWIP systems are systems where the number of items is kept constant. Here, the number of customers (or entities) is kept constant, which means that the next customer will only arrive when the current customer's service is completed. Because the number of customers is kept constant, the CONWIP system is expected to

have no uncertainty regarding $\widehat{Y_1}$ and, hence, the inputs X_1 and X_2 should have no impact on $\widehat{Y_1}$. Moreover, the customer's time in system is determined by how long it takes to be serviced plus time spent in the queue, if any. The goal of using the CONWIP system is to investigate the effectiveness of the entropy measures in capturing these known characteristics. Here, different from the previous experiments, the arrival process is determined by the service completion process.

Based on the aforementioned characteristics, one should expect the entropy of the average number of entities in the system, $\widehat{Y_1}$, to be zero. One would expect the entropy of the arrival process, X_1, the service time, X_2, and the average time in the system, $\widehat{Y_2}$ to be equal, as the arrival process and the time in system are dictated by the service time. However, in a simulation model two events, e.g., an arrival and service completion, cannot occur exactly at the same time. Therefore, some small differences should be expected in this case. Moreover, knowing that the output uncertainty may not be only comprised by the input uncertainty but also by some other uncertainties of the system (for instance, the computational limitation just previously mentioned), some small differences between X_2 and $\widehat{Y_2}$ are also expected.

From Figure 21, it is possible to observe that the entropy of $\widehat{Y_1}$ is only equal to zero for larger number of bins, i.e., number of bins greater than or equal to 1,000. For number of bins greater than or equal to 25, the entropy of $\widehat{Y_1}$ is close to 0, however it is not 0. For number of bins smaller than 25, the entropy of $\widehat{Y_1}$ is constant over the number of replications but not equal to 0, which means that regardless of the amount of data of the

simulation generated inputs X_1 and X_2 and the amount of data of the output $\widehat{Y_1}$, the measured uncertainty by the entropy method was the same but not null. This can be explained based on how the probability density is estimated using the histogram method. According to the histogram-based method, the probability density is estimated using $\hat{f}^{hist}{}_j(x) = \frac{1}{nh}\sum_{i=1}^{n} \mathbf{I}\{x_i \in [t_j, t_{j+1})\} for\ x_j \in B_j, j = 1, \dots, k$. Because in the CONWIP system the NIS is constant, all the $\widehat{Y_1}$ are equal and, hence, all $\widehat{Y_1}$ belong to the same *j* resulting in $\hat{f}^{hist}{}_j(x) = \frac{n}{nh}$. If h did not exist, then $\hat{f}^{hist}{}_j(x) = \frac{n}{n} = 1$ and $log\left(\hat{f}^{hist}{}_j(x)\right) = 0$. However, since h is fixed, $log\left(\hat{f}^{hist}{}_j(x)\right)$ is constant. The smaller the binwidth (or the larger the number of bins), the value of $log\left(\hat{f}^{hist}{}_j(x)\right)$ will be closer to 0. Similar analysis can be made for the mutual information. One would expect that the inputs X_1 and X_2 should have no impact on $\widehat{Y_1}$ and, hence, $I(X_1; Y_1)$ and $I(X_2; Y_1)$ should be equal to 0. However, from Figure 21 one can see that the MI is constant but not 0, despite being close to or equal to 0 for larger number of bins. The reason for this is that MI can be calculated by $I(X_1; Y_1) = H(X_1) + H(Y_1) - H(X_1, Y_1)$. In this case, $H(X_1, Y_1) = H(X_1)$. Hence, $I(X_1; Y_1) = H(Y_1)$.

From Figure 22 one can see that the entropy of X_2 and $\widehat{Y_2}$ are equal for number of bins equal to 2 and same number of replications, but they start to differ with the increase in the number of bins. The entropy of X_1 differs from the entropy of X_2 and $\widehat{Y_2}$ in every number of bins investigated.

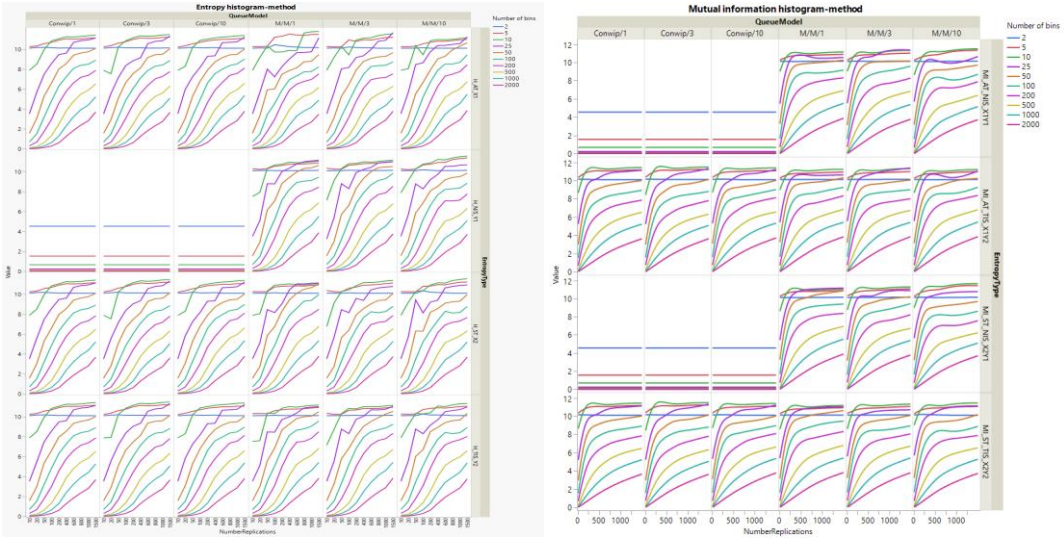

Figure 21. Entropy and MI measures per queue model per number of bins using histogram-based method with fixed number of bins and probability density function (CONWIP vs original experiments).

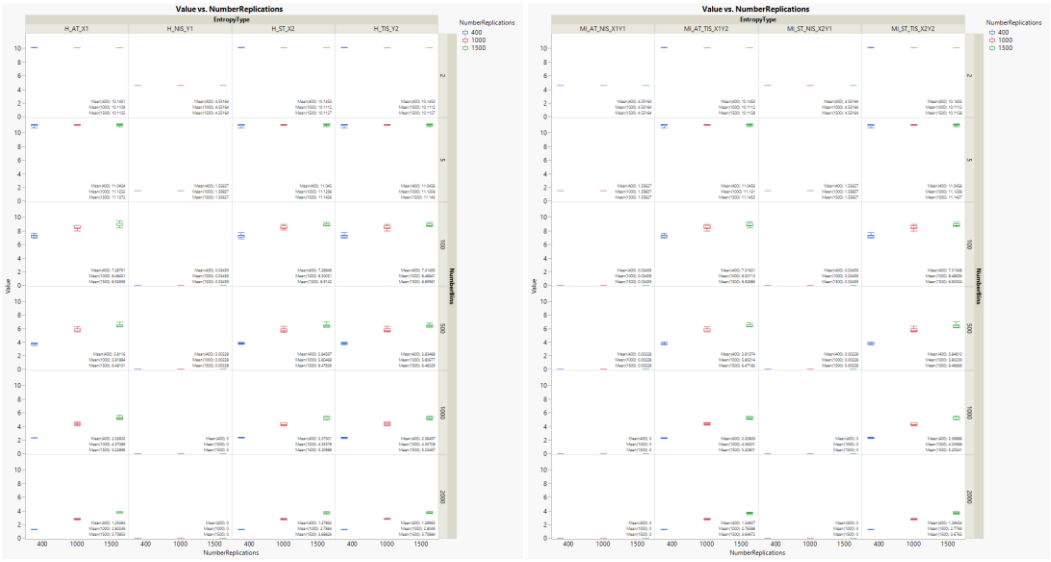

Figure 22. Entropy and MI measures per number of bins per replication using histogram-based method with fixed number of bins and probability density function for CONWIP systems.

Using the histogram-based method with fixed number of bins and discrete empirical distribution, the normalized entropy of $\widehat{Y_1}$, the NMI_{theor} between $\widehat{Y_1}$ and X_1, and the NMI_{theor} between $\widehat{Y_1}$ and X_2 are equal to zero regardless of the number of bins chosen, as shown in Figure 23. Similarly, using the histogram-based method with probability density function, the normalized entropy of X_1, X_2 and $\widehat{Y_2}$ are not equal regardless of the number of bins, which leads to similar results for the mutual information.

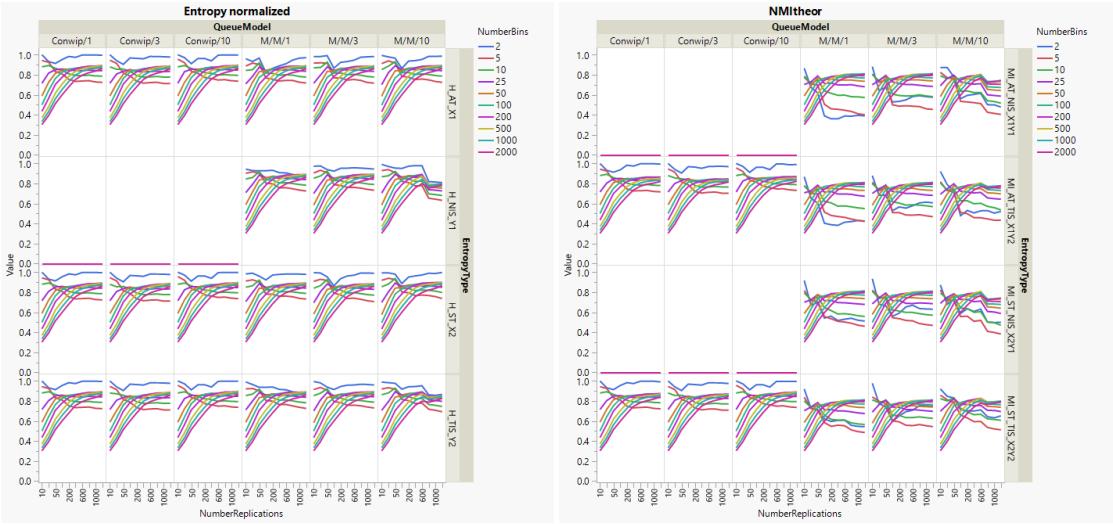

Figure 23. Normalized entropy and $\boldsymbol{NMI_{theor}}$ measures per queue model per number of bins using histogram-based method with fixed number of bins and discrete empirical distribution (CONWIP vs original experiments).

The NMI_{arith}, NMI_{geom}, and NMI_{joint} between $\widehat{Y_1}$ and X_1 and between $\widehat{Y_1}$ and X_2 are also equal to zero regardless of the number of bins chosen when using the histogram-

based method with fixed number of bins and discrete empirical distribution, as shown in Figure 24 and Figure 25. However, as already discussed, these normalized measures do not present a behavior similar to the measures calculated using the histogram-based method with fixed number of bins and probability density function. Instead, in this case the measures, when not equal to zero, tend to decrease with the increase in the number of replications. Because of that, NMI_{theor} appears to be a better normalization method overall.

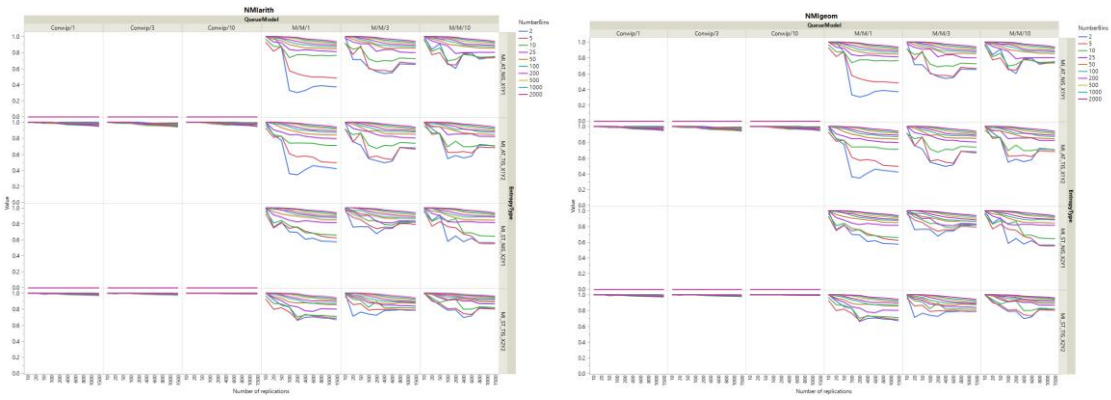

Figure 24. $\boldsymbol{NMI_{arith}}$ and $\boldsymbol{NMI_{geom}}$ measures per queue model per number of bins using histogram-based method with fixed number of bins and discrete empirical distribution (CONWIP vs original experiments).

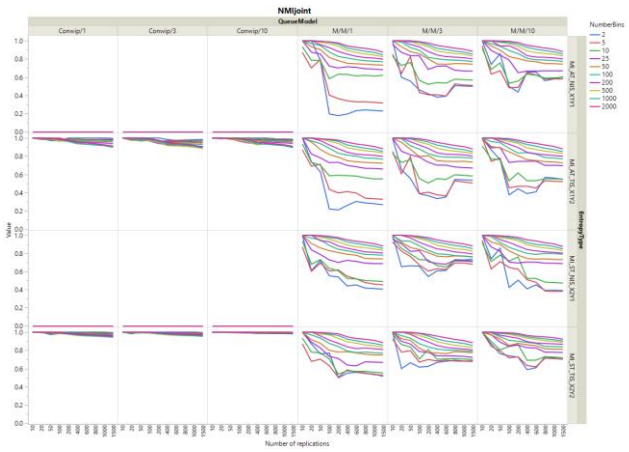

Figure 25. $\boldsymbol{NMI_{joint}}$ measures per queue model per number of bins using histogram-based method with fixed number of bins and discrete empirical distribution (CONWIP vs original experiments).

Next, a third input parameter X_3, namely travel time, was added to the M/M/s system. A total of 200 experiments with the third input was run. Experiments #931 to #1030 in Table 66 correspond to the initial experiments numbered #1 to #100, with the added third input as deterministic travel time of 10 minutes. Experiments #1031 to #1130 in Table 66 correspond to the initial experiments #1 to #100, with the added third input as stochastic travel time exponentially distributed with a mean of 10 minutes and using a fixed seed.

For the deterministic travel time input, one would expect the entropy of X_3 to be zero, as there is no uncertainty associated with the input. Similarly, one would expect this input to bring no reduction in the average uncertainty of the simulation outputs $\widehat{Y_1}$ and $\widehat{Y_2}$, which means that one would expect $\mathrm{I}(X_3; Y_1)$ and $\mathrm{I}(X_3; Y_2)$ to be equal to 0. For the

stochastic case, there is uncertainty associated with the travel time input and the entropy should capture it.

Using the histogram-based method with fixed number of bins and probability density function, the results obtained for the deterministic travel time are similar to the ones obtained for the average number in system in the CONWIP system using the same method. As shown in Figure 26, the entropy of X_3 is only equal to zero for larger number of bins, i.e., number of bins greater than or equal to 1,000. For number of bins greater than or equal to 25, the entropy of X_3 tends to go to 0, however it is not 0. For number of bins smaller than 25, the entropy of X_3 is constant over the number of replications but not equal to 0, which means that regardless of the amount of data of the simulation input X_3, the measured uncertainty by the entropy method was the same but not null. From Figure 26, a similar analysis can be done for $\mathrm{I}(X_3; Y_1)$ and $\mathrm{I}(X_3; Y_2)$. For the stochastic case, the entropy method captured the uncertainty of X_3 in a similar way that it did for the simulation generated inputs X_1 and X_2 and the simulation outputs $\widehat{Y_1}$ and $\widehat{Y_2}$.

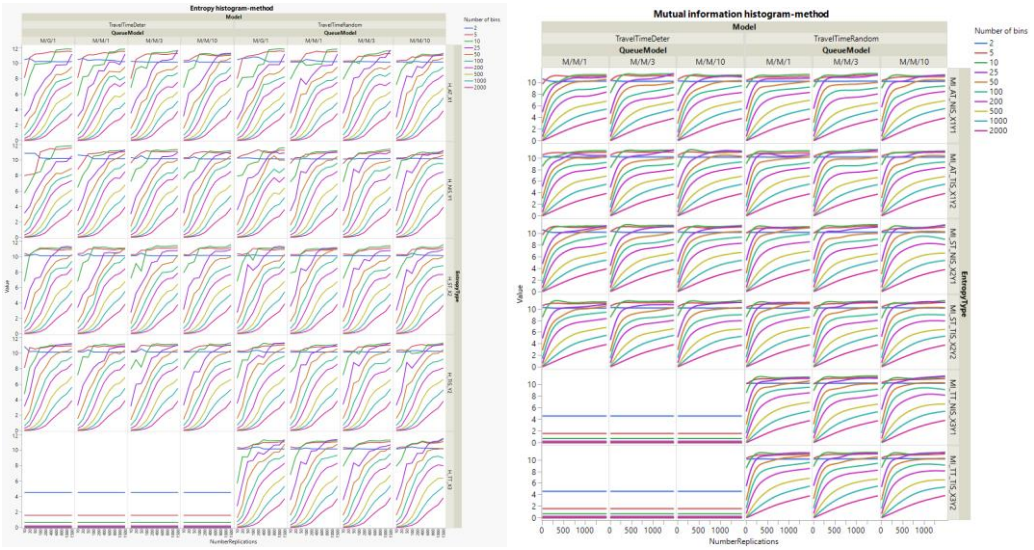

Figure 26. Entropy and MI measures per queue model per number of bins using histogram-based method with fixed number of bins and probability density function (deterministic travel time vs stochastic travel time).

Using the histogram-based method with fixed bins and discrete empirical distribution, the results for the deterministic travel time are also similar to the ones obtained for the average number in system in the CONWIP system using the same method. For the deterministic case, Figure 27 shows that the entropy of X_3 is equal to zero, as expected, because there is no information or uncertainty added into the system by X_3. Although the travel time is deterministic, the travel time input will still impact the output average number of entities in the system, $\hat{Y}_1$, and average time in the system, $\hat{Y}_2$. However, running experiments and getting information about X_3 does not provide any extra information about $\hat{Y}_1$ or $\hat{Y}_2$ nor reduce the uncertainty of these outputs, because running more experiments does not provide more information about X_3 than what was already known before running the experiments. Therefore, X_3 should impact the values of $\hat{Y}_1$ and

$\hat{Y}_2$, but the mutual information between X_3 and Y_1 and between X_3 and Y_2 are expected to be equal to zero and these are the results obtained, as shown in Figure 27. For the stochastic case, the entropy method captured the uncertainty of X_3 in a similar way that it did for the simulation generated inputs X_1 and X_2 and the simulation outputs $\widehat{Y_1}$ and $\widehat{Y_2}$.

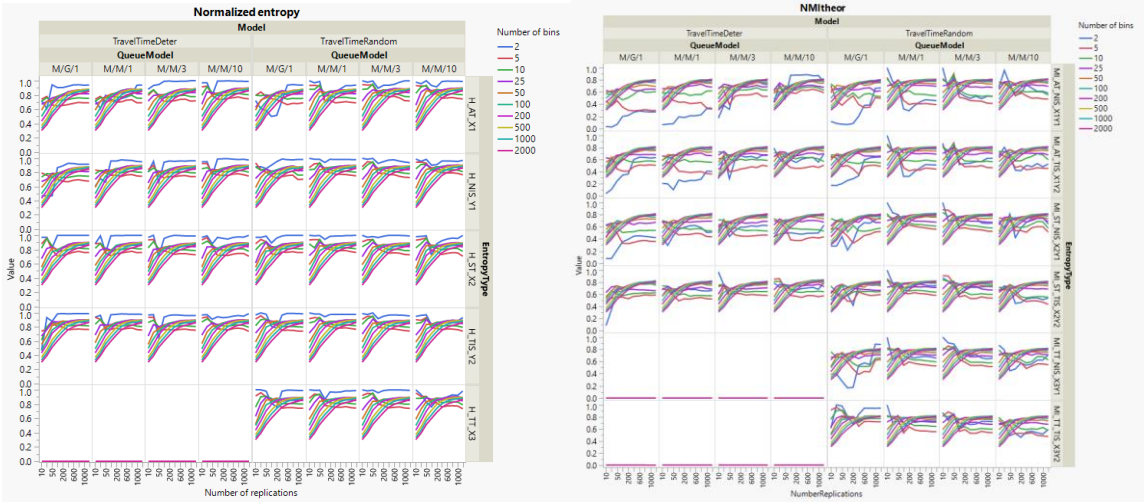

Figure 27. Normalized entropy and $\boldsymbol{NMI_{theor}}$ measures per queue model per number of bins using histogram-based method with fixed number of bins and discrete empirical distribution (deterministic travel time vs stochastic travel time).

2.4.4. Analysis of entropy and MI as a measure of uncertainty quantification in simulation models

Although the definition of uncertainty quantification is simple, developing a systematic method to quantify uncertainty and validating or assessing the potential of the proposed method is not an easy task. A queue model for which closed-form solutions are available was adopted in this dissertation as an attempt to assess the potential of the

measures as a method of uncertainty quantification in simulation models. As known in the simulation field, with the increase in the number of replications, one expects to reduce the intrinsic output-uncertainty. However, care must be taken because the extrinsic input-uncertainty may outweigh the intrinsic output-uncertainty and the total uncertainty may, thus, increase. But in general, one can say that the total uncertainty decreases with the increase in the number of replications.

The mutual information is the average reduction of the uncertainty of the output provided by the input and the entropy measure is the average information or uncertainty per input or output. These measures are typically measured in bits, alternatively called natural units (nats) or sometimes Shannons. The simulation generated outputs are usually measured in different units and in different magnitude of scale. For this reason, the Shannons unit cannot be directly compared to the simulation generated outputs or the theoretical outputs, which hinders the validation or the assessment of the potential of the proposed method.

Therefore, in order to investigate the potential of the entropy and MI as a method of uncertainty quantification in simulation model, the measures results were compared against results of methods commonly applied in the scientific community. For the entropy measures, the following comparisons were performed:

(i) The entropy measure (or average entropy measure) detects an increase or decrease in uncertainty with the increase in the number of replications that is in agreement with the detection by the error method being compared to.

(ii) The entropy measure (or average entropy measure) detects the experiment that leads to the maximum uncertainty in agreement with the detection by the error method being compared to.

For the comparisons, four error methods were considered: SAE, SSE, MAE, and MSE. For consistency in the comparison, the average entropy measure was considered instead of entropy measure when MAE or MSE was the error comparison method. This led to a total of eight comparisons per input or output per method used to calculate the entropy measures.

The reasons to perform the above comparisons are: (i) to understand whether the entropy measure agrees with other uncertainty methods of the literature; and, (ii) identify the experiment configuration with the highest uncertainty with respect to one of the inputs or outputs. This can help assessing the potential of the entropy measures as a method of uncertainty quantification in simulation models and determining the binwidth to use to investigate the impacts of input uncertainty on the outputs.

The methods SAE and SSE were calculated in R using Equation 44 and Equation 45, respectively.

$$SAE = \sum_{i=1}^{n} |y_i - \hat{y}_i| \qquad \text{Equation 44}$$

where y_i is the observed value and $\hat{y}_i$ is the predicted value. In this case, y_i and $\hat{y}_i$ can refer to either the inputs X_1, X_2 or X_3 or to the outputs Y_1 or Y_2. The predicted value is the value calculated through the $M/M/s$ or $M/G/1$ queue system exact solution and the observed value is the value resulting from the simulation experiment.

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad \text{Equation 45}$$

Average entropy is the entropy measure divided by the number of replications n and MAE and MSE are the SAE and SSE divided by the number of replications, respectively.

JMP® and Tableau® software were used to perform the comparisons and analysis. Each comparison was performed for the entropy measures calculated using: (1) the histogram-based method with fixed number of bins and probability density function; (2) the histogram-based method with optimum number of bins and probability density function; (3) the histogram-based method with fixed number of bins and discrete empirical distribution; and, (4) the histogram-based method with fixed number of bins, discrete empirical distribution and normalization of the entropy measure.

To calculate the increase or decrease in uncertainty with the increase in the number of replications, the entropy of one experiment was compared with another one with the same configuration and smaller number of replications, as given by Equation 46.

$$Change\ in\ uncertainty_{j+1} = \mathrm{H}(\mathrm{Z}_i)_{j+1} - \mathrm{H}(\mathrm{Z}_i)_j \; ; j = 1, \dots, 9 \qquad \text{Equation 46}$$

where Z_i is either the input X_1, X_2 or X_3 or the ouput Y_1 or Y_2 and j is the experiment. The experiments are divided into groups with equal configurations but different number of replications, which leads to groups of 10 experiments.

As shown in Table 6, considering the comparison to detect an increase or decrease in uncertainty the entropy measure calculated using fixed number of bins and probability

density function does not appear to have results in agreement with the SAE or SSE method for a low number of bins. This implies that, in general, less than 50% of the time the results of the entropy method matched the results of the SAE or SSE method for a number of bins with a value of 2. Interestingly, for lower number of bins, the entropy method shows better results with the outputs than the inputs. However, with the increase in the number of bins, the entropy measures for the inputs slightly outperform the entropy measures for the outputs. The results of the entropy method improve with an increase in the number of bins. From number of bins with a value of 50 or higher, one can see that the entropy method is in agreement with the SAE or SSE method more than 90% of the time. In fact, it is worth noting that for any of the methods used to calculate the entropy measures, the comparisons performed between the entropy measures and the SAE method led to exactly the same results as of the comparisons performed between the entropy measures and the SSE method.

Entropy measures are measures of total information of uncertainty in a system, similar to the SAE or SSE. It is clear that the entropy measures compared to SAE and SSE give better results than the average entropy measures compared to MAE and MSE. However, a possible issue with that comparison is that both methods (entropy measures for larger number of bins or SAE and SSE) have a tendency to increase with the increase in the number of replications, i.e., the amount of data. Hence, a better comparison that eliminates this possible bias towards the number of replications is to consider the MAE, MSE, and average entropy.

The average of the entropy measures calculated using histogram-based method, fixed number of bins, and probability density function was compared to MAE and MSE. In this case, it is observed that the performance of the method also improves with the increase in the number of bins. However, after reaching a point, the performance appears to reach its maxima and subsequently starts decreasing. For all, except for one comparison against MAE or MSE, the best performance of the entropy measure was obtained when the number of bins was 1,000, which indicates that this could potentially be a good number of bins to be chosen for this object of study (queue system). Even in the case where the number of bins equal to 1,000 did not give the best performance, it was able to give a performance very close to the best. Although the entropy measures when compared against the MAE or MSE perform poorer than when compared against SAE or SSE, if appropriate number of bins is chosen, the method can still achieve a performance of at least 60%.

Similar observations can be made in terms of the entropy measures performance to detect the experiment that leads to the maximum uncertainty. As shown in Table 7, the ability of the entropy measures in comparison to the SAE and SSE method increases with the increase in the number of bins. Similar behavior also occurs when the average entropy measures are compared to MAE or MSE. However, in the latter the performance decreases beyond a certain number of bins.

When the entropy measures are compared to SAE or SSE, there are, in general, different number of bins (e.g., from 500 to 2,000) for which the entropy measures show

the same performance. However, when the average entropy measures are compared to MAE or MSE it was observed that 1,000 bins usually lead to the maximum performance.

Table 6. Results from histogram-based method using fixed number of bins and probability density function for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Number of bins	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	2	35.67%	35.67%	44.30%	54.68%
	5	66.67%	66.67%	44.30%	54.68%
	10	80.99%	80.99%	44.30%	54.68%
	25	89.77%	89.77%	44.59%	54.97%
	50	95.32%	95.32%	50.15%	60.53%
	100	97.37%	97.37%	51.75%	62.43%
	200	97.37%	97.37%	43.42%	54.09%
	500	99.27%	99.27%	52.92%	62.72%
	1000	**100.00%**	**100.00%**	**61.55%**	**71.05%**
	2000	**100.00%**	**100.00%**	60.53%	60.67%
$H(X_2)$	2	40.94%	40.94%	45.39%	52.05%
	5	67.54%	67.54%	45.39%	52.05%
	10	77.78%	77.78%	45.39%	52.05%
	25	87.65%	87.65%	45.39%	52.05%
	50	96.13%	96.13%	49.34%	56.29%
	100	96.71%	96.71%	53.80%	61.33%
	200	98.98%	98.98%	47.30%	55.12%
	500	**100.00%**	**100.00%**	54.82%	63.82%
	1000	**100.00%**	**100.00%**	**63.60%**	**72.88%**
	2000	**100.00%**	**100.00%**	59.36%	65.42%
$H(X_3)$	2	75.00%	75.00%	19.44%	24.44%
	5	87.78%	87.78%	19.44%	24.44%
	10	88.89%	88.89%	19.44%	24.44%
	25	94.44%	94.44%	20.00%	25.00%
	50	99.44%	99.44%	21.67%	26.67%
	100	98.33%	98.33%	25.00%	30.00%

	200	98.33%	98.33%	21.67%	26.67%
	500	**100.00%**	**100.00%**	26.67%	31.67%
	1000	**100.00%**	**100.00%**	80.56%	**83.33%**
	2000	**100.00%**	**100.00%**	**81.11%**	81.67%
$H(Y_1)$	2	56.29%	56.29%	49.20%	54.90%
	5	74.93%	74.93%	48.76%	54.46%
	10	84.21%	84.21%	48.76%	54.46%
	25	87.87%	87.87%	48.90%	54.61%
	50	94.08%	94.08%	52.34%	57.89%
	100	95.91%	95.91%	56.43%	61.99%
	200	97.15%	97.15%	48.76%	55.19%
	500	99.12%	99.12%	54.24%	61.70%
	1000	**99.56%**	**99.56%**	**61.92%**	**68.49%**
	2000	**99.56%**	**99.56%**	56.51%	59.87%
$H(Y_2)$	2	53.87%	53.87%	48.10%	53.44%
	5	73.68%	73.68%	47.81%	53.14%
	10	82.53%	82.53%	47.81%	53.14%
	25	89.40%	89.40%	47.81%	53.14%
	50	95.32%	95.32%	51.68%	57.31%
	100	96.93%	96.93%	55.41%	61.77%
	200	97.73%	97.73%	49.12%	55.48%
	500	99.42%	99.42%	54.39%	61.62%
	1000	**99.71%**	**99.71%**	**61.99%**	**68.93%**
	2000	**99.71%**	**99.71%**	55.92%	61.11%

Table 7. Results from histogram-based method using fixed number of bins and probability density function for detecting the experiment that leads to the maximum uncertainty.

Entropy	Number of bins	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	2	0.00%	0.00%	10.53%	10.53%
	5	53.95%	53.95%	10.53%	10.53%
	10	85.53%	85.53%	10.53%	10.53%
	25	94.74%	94.74%	10.53%	10.53%
	50	100.00%	100.00%	9.21%	9.21%
	100	94.74%	94.74%	0.00%	0.00%

	200	98.68%	98.68%	0.00%	0.00%
	500	**100.00%**	**100.00%**	0.00%	0.00%
	1000	**100.00%**	**100.00%**	**75.00%**	**75.00%**
	2000	**100.00%**	**100.00%**	3.95%	3.95%
$H(X_2)$	2	3.95%	3.95%	2.63%	0.00%
	5	80.26%	80.26%	2.63%	0.00%
	10	85.53%	85.53%	2.63%	0.00%
	25	88.16%	88.16%	2.63%	0.00%
	50	94.08%	94.08%	0.00%	0.00%
	100	93.42%	93.42%	0.00%	0.00%
	200	97.37%	97.37%	3.95%	0.00%
	500	**100.00%**	**100.00%**	6.58%	10.53%
	1000	**100.00%**	**100.00%**	**69.08%**	**69.74%**
	2000	**100.00%**	**100.00%**	15.79%	15.13%
$H(X_3)$	2	50.00%	50.00%	50.00%	50.00%
	5	95.00%	95.00%	50.00%	50.00%
	10	90.00%	90.00%	50.00%	50.00%
	25	**100.00%**	**100.00%**	55.00%	50.00%
	50	**100.00%**	**100.00%**	55.00%	50.00%
	100	85.00%	85.00%	55.00%	50.00%
	200	85.00%	85.00%	50.00%	50.00%
	500	50.00%	50.00%	50.00%	55.00%
	1000	**100.00%**	**100.00%**	**75.00%**	**75.00%**
	2000	**100.00%**	**100.00%**	50.00%	50.00%
$H(Y_1)$	2	39.47%	39.47%	11.84%	9.87%
	5	76.32%	76.32%	11.84%	9.87%
	10	84.21%	84.21%	11.84%	9.87%
	25	71.71%	71.71%	11.84%	9.87%
	50	81.58%	81.58%	7.24%	9.21%
	100	88.16%	88.16%	3.29%	7.24%
	200	94.08%	94.08%	0.00%	1.32%
	500	**96.05%**	**96.05%**	2.63%	2.63%
	1000	**96.05%**	**96.05%**	**65.13%**	**61.84%**
	2000	**96.05%**	**96.05%**	11.84%	11.84%
$H(Y_2)$	2	28.95%	28.95%	9.21%	4.61%
	5	84.21%	84.21%	9.21%	4.61%
	10	84.21%	84.21%	9.21%	4.61%
	25	80.92%	80.92%	9.21%	4.61%
	50	92.76%	92.76%	8.55%	5.26%

100	89.47%	89.47%	1.97%	6.58%
200	94.74%	94.74%	2.63%	1.32%
500	97.37%	97.37%	5.26%	10.53%
1000	97.37%	97.37%	**57.89%**	**60.53%**
2000	**98.68%**	**98.68%**	14.47%	13.16%

For the histogram-based method using optimum number of bins, the entropy measures did not show good results in comparison to the SSE, SAE, MAE, and MSE methods for either detecting an increase or decrease in uncertainty with the increase in the number of replications or for detecting the experiment that leads to the maximum uncertainty, as shown in Table 8 and Table 9, respectively. Based on the results discussed for the histogram-based method using fixed number of bins, these poor results were expected because for the experiments in this study the number of bins turned out to be small when using the optimum number of bins rules (e.g., $\approx 10\ to\ 20$). Therefore, the results for the optimum number of bins are in accordance with the previous results for the fixed number of bins.

With the exception of the travel time input for detecting an increase or decrease in the uncertainty with the increase in the number of replications, the different optimum number of bins rules led to same results of performance among themselves when compared to MAE or MSE. When compared to SAE or SSE, Sturges' rule led to better performance for both detecting an increase or decrease in uncertainty and for detecting the experiment that leads to the maximum uncertainty. The only exception was again the travel input for detecting an increase or decrease in the uncertainty with the increase in the

number of replications. In this case, FD had a better performance. This is due to the fact that the travel input involved both stochastic and deterministic parameters, which affected how the number of bins were calculated for Scott's and FD's rules. The sample standard deviation and the interquartile range are taken into consideration in Scott's and FD's rules, respectively. When the data is deterministic, this would lead to an optimum binwidth of 0. To overcome this issue, only Sturges' rule was considered for the deterministic travel time input in this work. Table 10 shows that when one considers only the stochastic travel time, this input is not an exception anymore. The entropy measures calculated using the histogram-based method and optimum number of bins showed better performance when compared to SAE and SSE than when compared to MAE and MSE, similar to what was observed with fixed number of bins.

Table 8. Results from histogram-based method using optimum number of bins and probability density function for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Optimum number of bins rule	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	FD	59.06%	59.06%	**44.30%**	**54.68%**
	Scott	59.94%	59.94%	**44.30%**	**54.68%**
	Sturges	**73.10%**	**73.10%**	**44.30%**	**54.68%**
$H(X_2)$	FD	56.36%	56.36%	**45.39%**	**52.05%**
	Scott	61.99%	61.99%	**45.39%**	**52.05%**
	Sturges	**75.00%**	**75.00%**	**45.39%**	**52.05%**
$H(X_3)$	FD	**42.22%**	**42.22%**	**19.44%**	**24.44%**
	Scott	28.44%	28.44%	7.41%	8.33%
	Sturges	22.22%	22.22%	0.00%	0.00%
$H(Y_1)$	FD	56.65%	56.65%	**48.76%**	**54.46%**

	Scott	62.57%	62.57%	**48.76%**	**54.46%**
	Sturges	**74.85%**	**74.85%**	**48.76%**	**54.46%**
$H(Y_2)$	FD	56.94%	56.94%	**47.81%**	**53.14%**
	Scott	63.45%	63.45%	**47.81%**	**53.14%**
	Sturges	**75.44%**	**75.44%**	**47.81%**	**53.14%**

Table 9. Results from histogram-based method using optimum number of bins and probability density function for detecting the experiment that leads to the maximum uncertainty.

Entropy	Optimum number of bins rule	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	FD	42.11%	42.11%	**10.53%**	**10.53%**
	Scott	44.74%	44.74%	**10.53%**	**10.53%**
	Sturges	**69.74%**	**69.74%**	**10.53%**	**10.53%**
$H(X_2)$	FD	30.26%	30.26%	**2.63%**	**0.00%**
	Scott	46.71%	46.71%	**2.63%**	**0.00%**
	Sturges	**85.53%**	**85.53%**	**2.63%**	**0.00%**
$H(X_3)$	FD	75.00%	75.00%	**50.00%**	**50.00%**
	Scott	50.00%	50.00%	**50.00%**	**50.00%**
	Sturges	**90.00%**	**90.00%**	**50.00%**	**50.00%**
$H(Y_1)$	FD	26.32%	26.32%	**11.84%**	**9.87%**
	Scott	39.47%	39.47%	**11.84%**	**9.87%**
	Sturges	**75.66%**	**75.66%**	**11.84%**	**9.87%**
$H(Y_2)$	FD	32.89%	32.89%	**9.21%**	**4.61%**
	Scott	37.50%	37.50%	**9.21%**	**4.61%**
	Sturges	**80.26%**	**80.26%**	**9.21%**	**4.61%**

Table 10. Results for stochastic travel time only from histogram-based method using optimum number of bins and probability density function for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Optimum number of bins rule	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
	FD	62.22%	62.22%	**38.89%**	**48.89%**

H(X3) – TT (stochastic)	Scott	44.44%	44.44%	**38.89%**	**48.89%**
	Sturge	**74.44%**	**74.44%**	**38.89%**	**48.89%**

Finally, for the histogram-based method using fixed number of bins and discrete empirical distribution, observations similar to the histogram-based method with fixed number of bins and probability density function can be made for the comparison to SAE and SSE methods, as shown in Table 11 and

Table 12. The ability of the entropy measures in detecting an increase or decrease in uncertainty with the increase in the number of replications or in detecting the experiment that leads to the maximum uncertainty increases with the increase in the number of bins. The ability of the entropy measures appears to be greater for detecting the change in uncertainty than for detecting the scenario with the maximum uncertainty, as shown in Table 11 and

Table 12.

While the histogram-based method with probability density function detected change in uncertainty in agreement with the SAE or SSE method at least 90% of the time when number of bins is 50 or higher. For the histogram-based method with discrete empirical distribution this only occurs when number of bins is 200 or higher. The number of bins that led to performance greater than 90% in detecting the experiment with the

maximum uncertainty is also higher for the histogram-based method with discrete empirical distribution than for the histogram-based method with probability density function. As previously highlighted, the comparisons performed between the entropy measures and the SAE method led to exactly the same results as of the comparisons performed between the entropy measures and the SSE method.

In general, the entropy measures calculated using histogram-based method with fixed number of bins and discrete empirical distribution appear to have a better performance with the number of bins between 1,000 and 2,000. The entropy measures calculated using the histogram-based method, fixed number of bins and discrete empirical distribution show better performance when compared to SAE and SSE than when compared to MAE and MSE. When compared to MAE or MSE, the measures appear to perform worse than the entropy measures calculated using probability density function. From Table 11 and

Table 12 one can see that the number of bins has no impact on the performance of the average entropy measures when they are compared to MAE or MSE, different than the histogram-method with probability density function where this was only observed for lower number of bins. The average entropy measure calculated by the histogram-based method with discrete empirical distribution is monotonically increasing with the number of bins while the MAE or MSE are constant. Hence, the performance in detecting an increase or decrease in uncertainty or the maximum uncertainty becomes constant.

Results of the normalized entropy measures using the histogram-based method with fixed number of bins and discrete empirical distribution are not shown here because they are identical to the non-normalized results.

Figure 28 shows the results of the comparison for detecting an increase or decrease in uncertainty with the increase in the number of replications for $H(X_1)$ using the histogram-based method with fixed number of bins. From Figure 28, it is possible to compare the results from the histogram-based method with probability density function against the results from the histogram-based method with discrete empirical distribution.

Table 11. Results from histogram-based method using fixed number of bins and discrete empirical distribution for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Number of bins	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	2	59.21%	59.21%	44.30%	54.68%
	5	36.26%	36.26%	44.30%	54.68%
	10	34.50%	34.50%	44.30%	54.68%
	25	53.36%	53.36%	44.30%	54.68%
	50	63.01%	63.01%	44.30%	54.68%
	100	79.09%	79.09%	44.30%	54.68%
	200	96.05%	96.05%	44.30%	54.68%
	500	99.27%	99.27%	44.30%	54.68%
	1000	99.56%	99.56%	**44.30%**	**54.68%**
	2000	**100.00%**	**100.00%**	**44.30%**	**54.68%**
$H(X_2)$	2	54.75%	54.75%	45.39%	52.05%
	5	33.55%	33.55%	45.39%	52.05%
	10	38.52%	38.52%	45.39%	52.05%
	25	50.80%	50.80%	45.39%	52.05%
	50	65.42%	65.42%	45.39%	52.05%

	100	81.65%	81.65%	45.39%	52.05%
	200	95.39%	95.39%	45.39%	52.05%
	500	99.27%	99.27%	45.39%	52.05%
	1000	**100.00%**	**100.00%**	**45.39%**	**52.05%**
	2000	**100.00%**	**100.00%**	**45.39%**	**52.05%**
$H(X_3)$	2	74.44%	74.44%	69.44%	74.44%
	5	66.11%	66.11%	69.44%	74.44%
	10	71.11%	71.11%	69.44%	74.44%
	25	76.67%	76.67%	69.44%	74.44%
	50	85.00%	85.00%	69.44%	74.44%
	100	90.00%	90.00%	69.44%	74.44%
	200	98.33%	98.33%	69.44%	74.44%
	500	98.33%	98.33%	69.44%	74.44%
	1000	**100.00%**	**100.00%**	**69.44%**	**74.44%**
	2000	**100.00%**	**100.00%**	**69.44%**	**74.44%**
$H(Y_1)$	2	39.33%	39.33%	48.76%	54.46%
	5	24.93%	24.93%	48.76%	54.46%
	10	33.55%	33.55%	48.76%	54.46%
	25	49.12%	49.12%	48.76%	54.46%
	50	66.15%	66.15%	48.76%	54.46%
	100	76.17%	76.17%	48.76%	54.46%
	200	92.91%	92.91%	48.76%	54.46%
	500	97.51%	97.51%	48.76%	54.46%
	1000	**98.68%**	**98.68%**	**48.76%**	**54.46%**
	2000	**98.68%**	**98.68%**	**48.76%**	**54.46%**
$H(Y_2)$	2	40.94%	40.94%	47.81%	53.14%
	5	25.58%	25.58%	47.81%	53.14%
	10	32.38%	32.38%	47.81%	53.14%
	25	48.54%	48.54%	47.81%	53.14%
	50	65.06%	65.06%	47.81%	53.14%
	100	79.61%	79.61%	47.81%	53.14%
	200	94.74%	94.74%	47.81%	53.14%
	500	98.68%	98.68%	47.81%	53.14%
	1000	**99.42%**	**99.42%**	**47.81%**	**53.14%**
	2000	**99.42%**	**99.42%**	**47.81%**	**53.14%**

Table 12. Results from histogram-based method using fixed number of bins and discrete empirical distribution for detecting the experiment that leads to the maximum uncertainty.

Entropy	Number of bins	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	2	7.89%	7.89%	10.53%	10.53%
	5	0.00%	0.00%	10.53%	10.53%
	10	0.00%	0.00%	10.53%	10.53%
	25	0.00%	0.00%	10.53%	10.53%
	50	1.32%	1.32%	10.53%	10.53%
	100	7.89%	7.89%	10.53%	10.53%
	200	84.21%	84.21%	10.53%	10.53%
	500	**100.00%**	**100.00%**	**10.53%**	**10.53%**
	1000	**100.00%**	**100.00%**	**10.53%**	**10.53%**
	2000	**100.00%**	**100.00%**	**10.53%**	**10.53%**
$H(X_2)$	2	7.89%	7.89%	2.63%	0.00%
	5	0.00%	0.00%	2.63%	0.00%
	10	0.00%	0.00%	2.63%	0.00%
	25	0.00%	0.00%	2.63%	0.00%
	50	3.95%	3.95%	2.63%	0.00%
	100	26.32%	26.32%	2.63%	0.00%
	200	79.61%	79.61%	2.63%	0.00%
	500	98.68%	98.68%	2.63%	0.00%
	1000	**100.00%**	**100.00%**	**2.63%**	**0.00%**
	2000	**100.00%**	**100.00%**	**2.63%**	**0.00%**
$H(X_3)$	2	50.00%	50.00%	50.00%	50.00%
	5	50.00%	50.00%	50.00%	50.00%
	10	50.00%	50.00%	50.00%	50.00%
	25	50.00%	50.00%	50.00%	50.00%
	50	55.00%	55.00%	50.00%	50.00%
	100	60.00%	60.00%	50.00%	50.00%
	200	85.00%	85.00%	50.00%	50.00%
	500	85.00%	85.00%	50.00%	50.00%
	1000	**100.00%**	**100.00%**	**50.00%**	**50.00%**
	2000	**100.00%**	**100.00%**	**50.00%**	**50.00%**
$H(Y_1)$	2	0.00%	0.00%	11.84%	9.87%
	5	0.00%	0.00%	11.84%	9.87%
	10	0.00%	0.00%	11.84%	9.87%
	25	0.00%	0.00%	11.84%	9.87%

	50	1.32%	1.32%	11.84%	9.87%
	100	13.16%	13.16%	11.84%	9.87%
	200	68.42%	68.42%	11.84%	9.87%
	500	94.74%	94.74%	11.84%	9.87%
	1000	**96.05%**	**96.05%**	**11.84%**	**9.87%**
	2000	**96.05%**	**96.05%**	**11.84%**	**9.87%**
$H(Y_2)$	2	1.32%	1.32%	9.21%	4.61%
	5	0.00%	0.00%	9.21%	4.61%
	10	0.00%	0.00%	9.21%	4.61%
	25	0.00%	0.00%	9.21%	4.61%
	50	1.32%	1.32%	9.21%	4.61%
	100	19.74%	19.74%	9.21%	4.61%
	200	73.68%	73.68%	9.21%	4.61%
	500	93.42%	93.42%	9.21%	4.61%
	1000	**97.37%**	**97.37%**	**9.21%**	**4.61%**
	2000	**97.37%**	**97.37%**	**9.21%**	**4.61%**

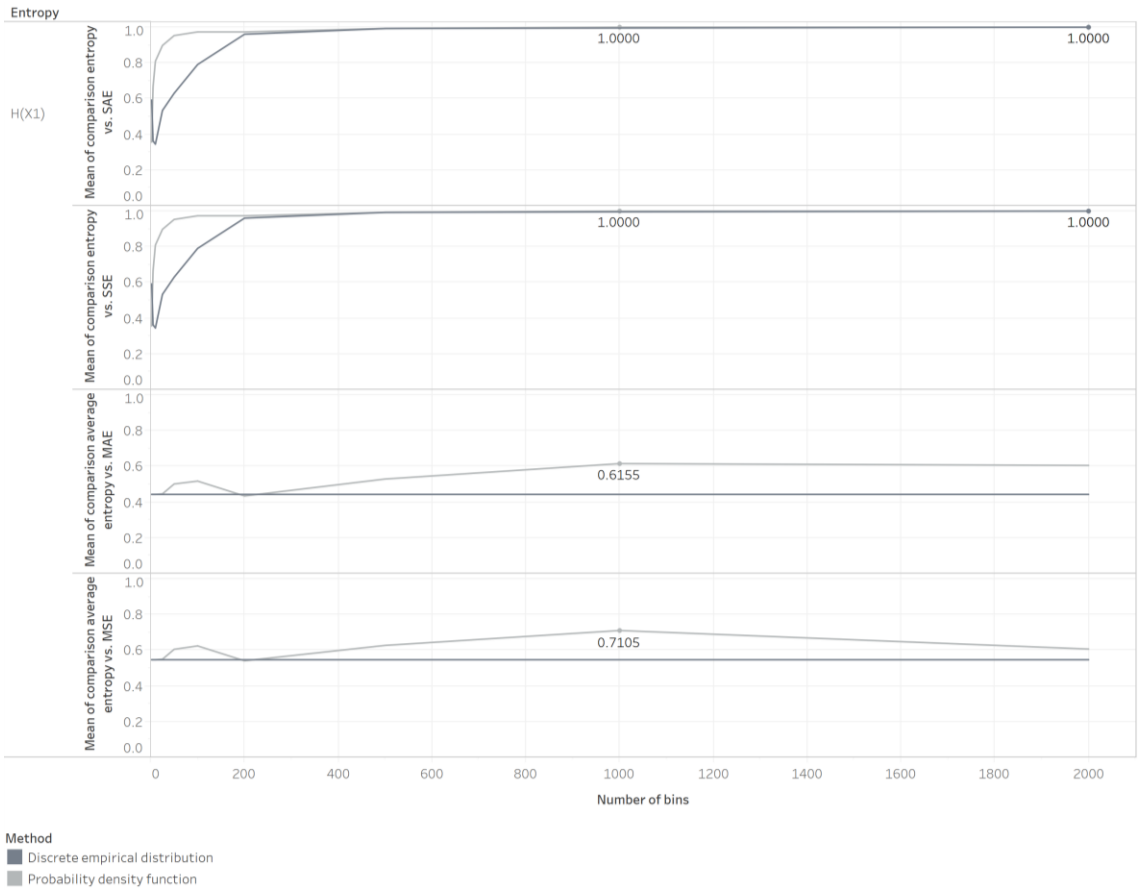

Figure 28. Results for detecting an increase or decrease in uncertainty with the increase in the number of replications for $\boldsymbol{H(X_1)}$ using the histogram-based method with fixed number of bins.

A final analysis performed in terms of the entropy measures was to compare the results from the histogram-based method with fixed number of bins and probability density function versus the results from the histogram-based method with fixed number of bins and discrete empirical distribution. The first is the theoretically correct approach but it brings challenges in its application. The second has practical benefits in terms of calculations and it is frequently used by practitioners, but it may not yield real benefits in

terms of applications. To compare the two methods, χ^2 test at an α-level of 0.05 was applied using JMP®, where the null hypothesis H_0 was that there is no difference between the results with the histogram-based method with probability density function and the results with the histogram-based method with discrete empirical distribution.

The results of the χ^2 test are shown in Table 13 to Table 18. As the results in Table 6, Table 7, Table 11, and Table 12 already indicated, Table 13 to Table 18 show that except for large number of bins, the results of the methods are actually statistically significantly different and, overall, the histogram-based method with probability density function shows better results than the histogram-based method with discrete empirical distribution for both detecting change in uncertainty and for detecting the experiment that leads to the maximum uncertainty in agreement with the SSE method (or the SAE method).

When the χ^2 test was performed on the average entropy measures in comparison to the MAE or MSE methods, the results were slightly different. The histogram-based method with probability density function was statistically significantly better than the histogram-based method with discrete empirical distribution only for higher number of bins (between 1,000 and 2,000). For a few cases where the number of bins was lower than 1,000, the histogram-based method with discrete empirical distribution was statistically significantly better than the histogram-based method with probability density function for detecting the experiment that leads to the maximum uncertainty.

Although there is not a consistency in the results, one can see that, in general, when the methods are statistically significantly different, the histogram-based method with

probability density function shows better results than the histogram-based method with discrete empirical distribution. However, this difference varies based on the methods the measures are being compared to and also on the number of bins used to calculate the measures.

Table 13. χ^2 test comparing the results of the histogram-based method using fixed number of bins and probability density function versus the results of the histogram-based method using fixed number of bins and discrete empirical distribution with respect to their capability of detecting an increase or decrease in uncertainty with the increase in the number of replications in agreement with the SSE method.

Proportion difference per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	-0.225	0.318	0.473	0.367	0.321	0.181	0.013	0.000	0.004	0.000
$H(X_2)$	-0.182	0.352	0.404	0.372	0.343	0.181	0.035	0.007	0.000	0.000
$H(X_3)$	0.006	0.217	0.178	0.036	0.144	0.083	0.000	0.017	0.000	0.000
$H(Y_1)$	0.163	0.481	0.487	0.373	0.269	0.190	0.041	0.016	0.008	0.008
$H(Y_2)$	0.126	0.488	0.509	0.410	0.301	0.171	0.029	0.007	0.003	0.003
P-value per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	0.1616	1.0000	0.0830	1.0000
$H(X_2)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0251**	1.0000	1.0000
$H(X_3)$	0.9038	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0007**	1.0000	0.0833	1.0000	1.0000
$H(Y_1)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0003**	**0.0194**	**0.0411**	**0.0411**
$H(Y_2)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0030**	0.1638	0.4136	0.4136

Table 14. χ^2 test comparing the results of the histogram-based method using fixed number of bins and probability density function versus the results of the histogram-based method using fixed number of bins and discrete empirical distribution with respect to

their capability of detecting the experiment that leads to the maximum uncertainty in agreement with the SSE method.

Proportion difference per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	-0.076	0.544	0.861	0.942	0.984	0.839	0.139	0.000	0.000	0.000
$H(X_2)$	-0.038	0.786	0.861	0.879	0.902	0.649	0.171	0.013	0.000	0.000
$H(X_3)$	0.000	0.450	0.400	0.500	0.450	0.250	0.000	-0.350	0.000	0.000
$H(Y_1)$	0.380	0.734	0.810	0.690	0.772	0.722	0.247	-0.025	0.000	0.000
$H(Y_2)$	0.266	0.824	0.848	0.809	0.915	0.674	0.202	0.038	0.000	0.013
P-value per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	**0.0134**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0015**	1.0000	1.0000	1.0000
$H(X_2)$	0.3062	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0006**	0.3189	1.0000	1.0000
$H(X_3)$	1.0000	**0.0013**	**0.0054**	**0.0003**	**0.0009**	0.0809	1.0000	**0.0181**	1.0000	1.0000
$H(Y_1)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	0.5165	1.0000	1.0000
$H(Y_2)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0004**	0.2492	1.0000	0.5630

Table 15. χ^2 test comparing the results of the histogram-based method using fixed number of bins and probability density function versus the results of the histogram-based method using fixed number of bins and discrete empirical distribution with respect to their capability of detecting an increase or decrease in uncertainty with the increase in the number of replications in agreement with the MAE method.

Proportion difference per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	0.000	0.000	0.000	0.003	0.058	0.075	0.009	0.086	0.173	0.162
$H(X_2)$	0.000	0.000	0.000	0.000	0.039	0.084	0.019	0.094	0.182	0.140
$H(X_3)$	0.500	0.500	0.500	0.494	0.478	0.444	0.478	0.428	0.111	0.117
$H(Y_1)$	0.004	0.000	0.000	0.001	0.036	0.077	0.000	0.055	0.132	0.077
$H(Y_2)$	0.003	0.000	0.000	0.000	0.039	0.076	0.013	0.066	0.142	0.081
P-value per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	1.0000	1.0000	1.0000	0.9134	**0.0303**	**0.0058**	0.7440	**0.0014**	**<0.0001**	**<0.0001**
$H(X_2)$	1.0000	1.0000	1.0000	1.0000	0.1425	**0.0018**	0.4796	**0.0004**	**<0.0001**	**<0.0001**

$H(X_3)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0149**	**0.0102**
$H(Y_1)$	0.8711	1.0000	1.0000	0.9569	0.1850	**0.0044**	1.0000	**0.0423**	**<0.0001**	**0.0041**
$H(Y_2)$	0.9139	1.0000	1.0000	1.0000	0.1519	**0.0048**	0.6266	**0.0148**	**<0.0001**	**0.0026**

Table 16. χ^2 test comparing the results of the histogram-based method using fixed number of bins and probability density function versus the results of the histogram-based method using fixed number of bins and discrete empirical distribution with respect to their capability of detecting an increase or decrease in uncertainty with the increase in the number of replications in agreement with the MSE method.

Proportion difference per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	0.000	0.000	0.000	0.003	0.058	0.077	0.006	0.080	0.164	0.060
$H(X_2)$	0.000	0.000	0.000	0.000	0.042	0.093	0.031	0.118	0.208	0.134
$H(X_3)$	0.500	0.500	0.500	0.494	0.478	0.444	0.478	0.428	0.089	0.072
$H(Y_1)$	0.004	0.000	0.000	0.001	0.034	0.075	0.007	0.072	0.140	0.054
$H(Y_2)$	0.003	0.000	0.000	0.000	0.042	0.086	0.023	0.085	0.158	0.080
P-value per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	1.0000	1.0000	1.0000	0.9136	**0.0286**	**0.0036**	0.8283	**0.0025**	**<0.0001**	**0.0249**
$H(X_2)$	1.0000	1.0000	1.0000	1.0000	0.1152	**0.0005**	0.2545	**<0.0001**	**<0.0001**	**<0.0001**
$H(X_3)$	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**<0.0001**	**0.0389**	0.0984
$H(Y_1)$	0.8704	1.0000	1.0000	0.9566	0.1995	**0.0046**	0.7856	**0.0065**	**<0.0001**	**0.0428**
$H(Y_2)$	0.9137	1.0000	1.0000	1.0000	0.1213	**0.0012**	0.3852	**0.0015**	**<0.0001**	**0.0028**

Table 17. χ^2 test comparing the results of the histogram-based method using fixed number of bins and probability density function versus the results of the histogram-based method using fixed number of bins and discrete empirical distribution with respect to their capability of detecting the experiment that leads to the maximum uncertainty in agreement with the MAE method.

Proportion difference per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	0.000	0.000	0.000	0.000	-0.013	-0.105	-0.105	-0.105	0.645	-0.066

$H(X_2)$	0.000	0.000	0.000	0.000	-0.026	-0.026	0.013	0.039	0.664	0.132
$H(X_3)$	0.000	0.000	0.000	0.050	0.050	0.050	0.000	0.000	0.250	0.000
$H(Y_1)$	0.000	0.000	0.000	0.000	-0.046	-0.086	-0.118	-0.092	0.533	0.000
$H(Y_2)$	0.000	0.000	0.000	0.000	-0.007	-0.072	-0.066	-0.039	0.487	0.053
P-value per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	1.0000	1.0000	1.0000	1.0000	0.7873	**0.0035**	**0.0035**	**0.0035**	**<0.0001**	0.1191
$H(X_2)$	1.0000	1.0000	1.0000	1.0000	0.1566	0.1566	0.6519	0.2485	**<0.0001**	**0.0040**
$H(X_3)$	1.0000	1.0000	1.0000	0.7590	0.7590	0.7590	1.0000	1.0000	0.1077	1.0000
$H(Y_1)$	1.0000	1.0000	1.0000	1.0000	0.3324	**0.0440**	**0.0018**	**0.0285**	**<0.0001**	1.0000
$H(Y_2)$	1.0000	1.0000	1.0000	1.0000	0.8864	**0.0490**	0.0868	0.3509	**<0.0001**	0.3185

Table 18. χ^2 test comparing the results of the histogram-based method using fixed number of bins and probability density function versus the results of the histogram-based method using fixed number of bins and discrete empirical distribution with respect to their capability of detecting the experiment that leads to the maximum uncertainty in agreement with the MSE method.

Proportion difference per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	0.000	0.000	0.000	0.000	-0.013	-0.105	-0.105	-0.105	0.645	-0.066
$H(X_2)$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.105	0.697	0.151
$H(X_3)$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.250	0.000
$H(Y_1)$	0.000	0.000	0.000	0.000	-0.007	-0.026	-0.086	-0.072	0.520	0.020
$H(Y_2)$	0.000	0.000	0.000	0.000	0.007	0.020	-0.033	0.059	0.559	0.086
P-value per number of bins										
Entropy	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**	**1,000**	**2,000**
$H(X_1)$	1.0000	1.0000	1.0000	1.0000	0.7873	**0.0035**	**0.0035**	**0.0035**	**<0.0001**	0.1191
$H(X_2)$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	**0.0035**	**<0.0001**	**0.0003**
$H(X_3)$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.7590	0.1077	1.0000
$H(Y_1)$	1.0000	1.0000	1.0000	1.0000	0.8879	0.5565	**0.0196**	0.0622	**<0.0001**	0.6956
$H(Y_2)$	1.0000	1.0000	1.0000	1.0000	0.8500	0.5811	0.2206	0.1646	**<0.0001**	0.0618

In information theory, the mutual information between X and Y is the average reduction in uncertainty in the value of Y provided by the value of X and vice-versa.

Mutual information is also a measure of the dependence or association between the variables X and Y. Therefore, in order to investigate the potential of the MI as a method of uncertainty quantification in simulation model, the MI results were compared to the results of three other measures of dependence between variables: distance correlation, Pearson correlation, and ${R^2}_{adj}$. To assess the results, the following analyses were performed:

(i) The MI is capable of identifying the input X_i that has the greatest impact on the uncertainty of the output Y_j in agreement with the measure of dependence.

(ii) The MI is capable of identifying the input X_i that has the least impact on the uncertainty of the output Y_j in agreement with the measure of dependence.

where $i = 1,2$ or $i = 1,2,3$ depending on the scenario being evaluated, and $j = 1,2$.

Distance correlation was proposed by Székely et al. (2007) as a measure of dependence between two random variables that is able to capture both linear and non-linear association. The distance correlation of the inputs and outputs were calculated using the package "energy" in the software R. Pearson correlation and ${R^2}_{adj}$ were also calculated in R. The latter was calculated using Equation 47.

$${R^2}_{adj} = 1 - (1 - R^2)\frac{n-1}{n-p-1} \qquad \text{Equation 47}$$

where n is the sample size and p is the total number of independent variables.

In order to compare the MI with the measures of dependence, an analysis procedure was developed because there could be experiments where either the MI or the measures of dependence would identify more than one input as the one with the greatest impact on the output. Similarly, there could be experiments where the MI or the measures of dependence would identify more than one input as the one with the least impact on the output. For the experiments that could have more than one input selected by the measures, a procedure was needed to identify whether the MI and the measure of dependence results were consistent.

For experiments that involved 3 inputs, that is, where arrival time, service time, and travel time were involved, the following procedure was followed:

1. If the MI identified 3 inputs as having the greatest (or the least) impact on the output, it means that none of the inputs are different than the other in terms of the uncertainty impact on the output. In this case, if the measure of dependence identified less than 3 inputs as having the greatest (or the least) impact on the output, then the MI and the measure of dependence did not agree among themselves. That is, the results were not consistent. Similarly, if the measure of dependence identified 3 inputs as having the greatest (or the least) impact on the output and the MI identified less than 3 inputs as having the greatest (or the least) impact on the output, then the MI and the measure of dependence did not agree among themselves.

2. If the MI identified 3 inputs as having the greatest (or the least) impact on the output and the measure of dependence identified 3 inputs as having the greatest (on the least) impact on the output, then the MI and the measure of dependence strongly agree among themselves, as all the choices among them were exactly the same.
3. If the MI identified i inputs as having the greatest (or the least) impact on the output and the measure of dependence identified i inputs as having the greatest (or the least) impact on the output, where $i = 1,2$, then:
    a. If all the inputs identified by the MI are the same as the ones identified by the measure of dependence, then the measures strongly agree.
    b. If at least one input identified by the measures is the same, then there is a weak agreement between the measures (this case is not possible for $i = 1$).
    c. Otherwise, the measures do not agree.
4. If the MI identified i inputs as having the greatest (or the least) impact on the output and the measure of dependence identified $i - 1$ inputs as having the greatest (or the least) impact on the output, where $i = 2,3$, then:
    a. If at least one input identified by the measures is the same, then there is a weak agreement between the measures.
    b. Otherwise, the measures do not agree.

The procedure is summarized on Table 19. A similar procedure was implemented for experiments that involved only 2 inputs (arrival time and service time).

Table 19. Procedure to identify whether the MI and the measure of dependence agree or not.

Method / Number of inputs with impact on output		Measure of dependence		
		3	**2**	**1**
MI	**3**	Strongly agree	Do not agree	Do not agree
	2	Do not agree	If: - all inputs are equal: strongly agree - at least one input is equal: weakly agree	If: - at least one input is equal: weakly agree - otherwise: do not agree
	1	Do not agree	If: - at least one input is equal: weakly agree - otherwise: do not agree	If: - input is equal: strongly agree - otherwise: do not agree

The results of the comparisons of the MI versus the distance correlation, the Pearson correlation, and the $R^2{}_{adj}$ methods are shown in Table 67 to Table 76 of the Appendix. Table 67 and Table 68 show the results when the MI is calculated using the histogram-based method with fixed number of bins and probability density function.

Table 69 and Table 70 show the results when the MI is calculated using the histogram-based method with optimum number of bins and probability density function.

Table 71 to Table 76 show the results when the MI is calculated using the histogram-based method with fixed number of bins and discrete empirical distribution. For the results shown in Table 67 to Table 76, either a weak or strong agreement, as defined in the aforementioned procedure, were considered an agreement among the measures.

First, logistic regression with two-factor interaction effect was performed using JMP® to investigate which factors affected the performance of the MI when compared against the different measures of dependence. Next, χ^2 test was performed using JMP® to investigate whether the performance of MI was statistically significantly different based on the method used to calculate the MI, the dependence measure to which the MI was being compared to, or the output that was being investigated.

Logistic regression was performed on the results of MI compared with each measure of dependence (distance correlation, Pearson correlation, and ${R^2}_{adj}$), from each calculation method (fixed bins with probability density function, fixed bins with discrete empirical distribution, and optimum number of bins with probability density function), and from each combination of greatest or least impact on the simulation output (input with the greatest impact on the NIS, input with the least impact on the NIS, input with the greatest impact on the TIS, and input with the least impact on the TIS). For the logistic regression, the following factors were considered as possible dependent variables that could affect the performance of the MI (the independent variable): (i) the number of bins (2, 5, 10, 25, 50, 100, 200, 500, 1000, 2000), (ii) the number of replications (10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500), and (iii) the normalization method (non-normalization, arith, joint, geom, and theor). The normalization method was only considered as a factor when the calculation method was the discrete empirical distribution, as it was the only method where normalization between 0 and 1 was possible and also

where normalization was recommended by the literature due to the impact of the number of bins.

Based on the whole model test, which compares the whole-model fit to the model that omits all the logistic regression parameters except the intercepts, the model considered here was statistically a better fit than the intercepts (p-value less than 0.0001 at α-level of 0.05). This test is analogous to the ANOVA for continuous responses. The null hypothesis is that the model fits no better than the model that includes only the intercepts, while the alternative hypothesis is that the model fits better than the model that includes only the intercepts. Table 20 shows the misclassification rate of the logistic model by method of calculation, measure of dependence, and impact on simulation output. Considering only the aforementioned factors as variables that impact the performance of the MI led to an average misclassification rate of 36.8%.

Table 20. Misclassification rate of the logistic regression model by method of calculation, measure of dependence, and impact on simulation output.

Method of calculation	Measure of dependence	Impact on simulation output	Misclassification rate
Fixed bins - probability density function	Distance correlation	Greatest impact on NIS	0.3691
Fixed bins - probability density function	Distance correlation	Least impact on NIS	0.3996
Fixed bins - probability density function	Distance correlation	Greatest impact on TIS	0.3701
Fixed bins - probability density function	Distance correlation	Least impact on TIS	0.3999
Fixed bins - probability density function	Pearson correlation	Greatest impact on NIS	0.3647

Fixed bins - probability density function	Pearson correlation	Least impact on NIS	0.3738
Fixed bins - probability density function	Pearson correlation	Greatest impact on TIS	0.3715
Fixed bins - probability density function	Pearson correlation	Least impact on TIS	0.3910
Fixed bins - probability density function	$R^2{}_{adj}$	Greatest impact on NIS	0.3586
Fixed bins - probability density function	$R^2{}_{adj}$	Least impact on NIS	0.3689
Fixed bins - probability density function	$R^2{}_{adj}$	Greatest impact on TIS	0.3684
Fixed bins - probability density function	$R^2{}_{adj}$	Least impact on TIS	0.3796
Optimum number of bins - probability density function	Distance correlation	Greatest impact on NIS	0.4413
Optimum number of bins - probability density function	Distance correlation	Least impact on NIS	0.4363
Optimum number of bins - probability density function	Distance correlation	Greatest impact on TIS	0.3605
Optimum number of bins - probability density function	Distance correlation	Least impact on TIS	0.3265
Optimum number of bins - probability density function	Pearson correlation	Greatest impact on NIS	0.4024
Optimum number of bins - probability density function	Pearson correlation	Least impact on NIS	0.3820
Optimum number of bins - probability density function	Pearson correlation	Greatest impact on TIS	0.3336
Optimum number of bins - probability density function	Pearson correlation	Least impact on TIS	0.2994
Optimum number of bins - probability density function	$R^2{}_{adj}$	Greatest impact on NIS	0.4006
Optimum number of bins - probability density function	$R^2{}_{adj}$	Least impact on NIS	0.3870
Optimum number of bins - probability density function	$R^2{}_{adj}$	Greatest impact on TIS	0.3342
Optimum number of bins - probability density function	$R^2{}_{adj}$	Least impact on TIS	0.3000
Fixed bins - discrete empirical distribution	Distance correlation	Greatest impact on NIS	0.4051

Fixed bins - discrete empirical distribution	Distance correlation	Least impact on NIS	0.3991
Fixed bins - discrete empirical distribution	Distance correlation	Greatest impact on TIS	0.3752
Fixed bins - discrete empirical distribution	Distance correlation	Least impact on TIS	0.3460
Fixed bins - discrete empirical distribution	Pearson correlation	Greatest impact on NIS	0.3837
Fixed bins - discrete empirical distribution	Pearson correlation	Least impact on NIS	0.3430
Fixed bins - discrete empirical distribution	Pearson correlation	Greatest impact on TIS	0.3598
Fixed bins - discrete empirical distribution	Pearson correlation	Least impact on TIS	0.3248
Fixed bins - discrete empirical distribution	$R^2{}_{adj}$	Greatest impact on NIS	0.3799
Fixed bins - discrete empirical distribution	$R^2{}_{adj}$	Least impact on NIS	0.3453
Fixed bins - discrete empirical distribution	$R^2{}_{adj}$	Greatest impact on TIS	0.3578
Fixed bins - discrete empirical distribution	$R^2{}_{adj}$	Least impact on TIS	0.3197

Table 21 shows the p-value and order of importance for the factors and effects in the logistic regression model for an α-level of 0.05. As shown in Table 21, number of replications appears to be the most important factor. With a few exceptions when using the optimum number of bins rule, the number of replications was the first factor in order of importance regardless of the method of calculation. With a few exceptions also when using the optimum number of bins rule, number of bins was the third factor in order of importance regardless of the method of calculation. Despite being the third factor in order of importance in the model, number of bins was still statistically significantly important

for the model in the majority of the cases. The normalization method and its interaction effects were not statistically significantly for the model when using the discrete empirical distribution, with the exception when considering the greatest impact on NIS. However, for the normalization method, care must be taken. As showed in Figure 9, the different normalized versions of the MI have different behaviors. Therefore, it is possible that if only two methods were being investigated, the results would be different.

Table 21. P-value and order of importance of factors on logistic regression model.

Method of calculation	Measure of dependence	Factor / P-value (order of importance)	Greatest impact on NIS	Least impact on NIS	Greatest impact on TIS	Least impact on TIS
Fixed number of bins - probability density function	Distance correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	0.0725 (3)	**0.0122 (3)**	**0.0001 (3)**	**0.0000 (3)**
		Number of replications x number of bins	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**
	Pearson correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0019 (3)**	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**
		Number of replications x number of bins	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**
	R^2_{adj}	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0034 (3)**	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**
		Number of replications x number of bins	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**
Optimum number of bins - probability density function	Distance correlation	Number of replications	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0469 (3)**	0.7034 (3)	**0.0000 (3)**	**0.0040 (3)**
		Number of replications x number of bins	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (2)**
	Pearson correlation	Number of replications	**0.0000 (3)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0000 (2)**	**0.0000 (3)**	**0.0000 (2)**	**0.0002 (3)**

		Number of replications x number of bins	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (3)**	**0.0000 (2)**
	R^2_{adj}	Number of replications	**0.0000 (2)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (2)**	**0.0000 (3)**
		Number of replications x number of bins	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (3)**	**0.0000 (2)**
Fixed number of bins - discrete empirical distribution	Distance correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**
		Normalization method	0.9681 (4)	0.6725 (4)	0.9987 (4)	0.9981 (4)
		Number of replications x number of bins	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**
		Number of replications x normalization method	1.0000 (5)	1.0000 (5)	0.9995 (5)	0.9997 (5)
		Number of bins x normalization method	1.0000 (6)	1.0000 (6)	1.0000 (6)	1.0000 (6)
	Pearson correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**
		Normalization method	**0.0248 (4)**	0.1421 (4)	0.9241 (4)	0.9880 (4)
		Number of replications x number of bins	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**
		Number of replications x normalization method	0.9996 (5)	0.9989 (5)	0.9929 (5)	0.9966 (5)
		Number of bins x normalization method	0.9999 (6)	0.9999 (6)	1.0000 (6)	1.0000 (6)
	R^2_{adj}	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of bins	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (3)**
		Normalization method	**0.0444 (4)**	0.1502 (4)	0.9440 (4)	0.9988 (5)
		Number of replications x number of bins	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**
		Number of replications x normalization method	0.9993 (5)	0.9992 (5)	0.9922 (5)	0.9957 (4)
		Number of bins x normalization method	0.9999 (6)	0.9999 (6)	1.0000 (6)	1.0000 (6)

Table 67 to Table 76 of the Appendix show the results of the comparison in a granulated level, which makes it harder to perform the comparison among the different methods and different measures of dependence. Table 22 shows the results summarized by calculation method, normalization method, measure of dependence, and impact on the output, where the green color represents the highest value per group, yellow the median, and red the lowest value per group. Overall, distance correlation is the measure of dependence that leads to the best performance when compared to the MI calculated using fixed number of bins with probability density function, while ${R^2}_{adj}$ is the measure of dependence that leads to the worst performance. For MI calculated using optimum number of bins rule and using fixed number of bins with discrete empirical distribution, Pearson correlation is the measure of dependence that leads to the best performance, while distance correlation is the measure of dependence that leads to the worst performance. However, it is important noting that there is no consistency in these results as this varies based on the number of bins used to calculate the MI. Whenever fixed number of bins is used to calculate the MI with either probability density function or discrete empirical distribution, depending on the number of bins used, either Pearson or distance correlation can lead to the best performance of the measure. Also, these differences may not be statistically significantly different as it will be investigated in the χ^2 test later in this section. The results of the χ^2 test are shown in Table 26.

An interesting observation that can also be made from Table 22 is that the results of the MI and the NMI_{theor} are very similar, with differences only in the decimals, which

is another indication that NMI_{theor} is the best normalization method for the mutual information.

Table 22. Results from the comparison of the measures of dependence versus the MI summarized by calculation method, normalization method, measure of dependence, and impact on the output.

Calculation method	Normalization method	Measure of dependence	Greatest impact on NIS [%]	Least impact on NIS [%]	Greatest impact on TIS [%]	Least impact on TIS [%]	Total [%]
Probability density function	Non-normalization	Distance correlation	37.82	42.75	37.75	43.96	40.57
		Pearson correlation	37.48	40.42	37.95	43.16	39.75
		R^2_{adj}	36.87	39.70	37.63	42.03	39.06
Optimum number of bins rule	Non-normalization	Distance correlation	49.03	55.40	56.71	61.90	55.76
		Pearson correlation	56.37	61.35	60.08	64.68	60.62
		R^2_{adj}	56.03	60.55	59.75	63.54	59.97
Discrete empirical distribution	Non-normalization	Distance correlation	44.28	49.71	52.96	56.43	50.84
		Pearson correlation	49.71	53.58	54.65	57.14	53.77
		R^2_{adj}	48.94	52.81	54.34	56.00	53.02
	Arith	Distance correlation	44.59	50.48	52.68	56.16	50.98
		Pearson correlation	47.81	51.97	54.08	56.75	52.65
		R^2_{adj}	47.09	51.20	53.78	55.61	51.92
	Joint	Distance correlation	44.58	50.48	52.68	56.16	50.98
		Pearson correlation	47.81	51.97	54.09	56.75	52.66
		R^2_{adj}	47.08	51.20	53.80	55.61	51.92
	Geom	Distance correlation	44.61	50.48	52.65	56.15	50.97
		Pearson correlation	47.82	51.97	54.00	56.68	52.62
		R^2_{adj}	47.10	51.20	53.71	55.54	51.89
	Theor	Distance correlation	44.24	49.67	52.95	56.42	50.82

	Pearson correlation	49.68	53.56	54.65	57.14	53.76
	$R^2{}_{adj}$	48.91	52.78	54.34	56.00	53.01

Table 23 shows the results summarized by calculation method, number of bins, measure of dependence, and impact on the output, where the green color represents the highest value per measure of dependence within calculation method and red represents the lowest values per measure of dependence within calculation method. When the MI is calculated using fixed number of bins with probability density function there is not a clear conclusion regarding the impact of the number of bins on the performance of the MI. When NIS is the output being considered, 2 is the number of bins that leads to the worst performance; and 10, for distance correlation, or 50, for either $R^2{}_{adj}$ or Pearson correlation, are the number of bins that lead to the best performance. When TIS is the output being considered, 2 is still the number of bins that leads to the worst performance when distance correlation is the measure of dependence being compared to. However, values of 1000 or 2000 are the number of bins that lead to worst performance when $R^2{}_{adj}$ or Pearson correlation are the measures of dependence considered. For best performance, the numbers of bins are 5 or 25 for $R^2{}_{adj}$ or Pearson correlation and 50 or 100 for distance correlation. Results are found to be more consistent when the MI is calculated using fixed number of bins with discrete empirical distribution. Overall, 2 is the number of bins that leads to the worst performance regardless of the measure of dependence and the output and 5 is the number of bins that leads to the best performance. Although there is a lack of

consistency when using probability density function, by comparing the results with the discrete empirical distribution, one can see that it appears that either larger number of bins (between 1000 and 2000) or very low number of bins (around 2) leads to worst performance in terms of MI, while mid-range (between 10 to 100) leads to best performance. However, it is important noting that as the results of Table 21 show, there is a interaction between the number of replications and number of bins and, in general, the interaction was shown to be more important than the number of bins alone. Therefore, ideally one should not evaluate the MI performance by looking at the number of bins alone, which could justify the lack of consistency.

When the MI is calculated using the optimum number of bins rule, it is interesting to note that regardless of the measure of dependence and the output, Sturges' rule appears to be the one that leads to the worst performance, while Scott's rule leads to best performance. FD's rule leads to the best performance when compared to Pearson correlation or ${R^2}_{adj}$ when considering the least impact on the NIS, but even in those cases Scott's rule performance is very close to FD's rule.

Table 23. Results from the comparison of the measures of dependence versus the MI summarized by calculation method, number of bins, measure of dependence, and impact on the output considering non-normalized version only.

Calculation method	Number of bins	Measure of dependence	Greatest impact on NIS [%]	Least impact on NIS [%]	Greatest impact on TIS [%]	Least impact on TIS [%]
	2	Distance correlation	33.42	37.85	31.90	38.73

		Pearson correlation	32.66	34.81	34.18	39.49
		$R^2{}_{adj}$	32.15	34.05	33.80	38.35
		2 Total	**32.74**	**35.57**	**33.29**	**38.86**
	5	Distance correlation	36.08	42.66	38.73	44.05
		Pearson correlation	36.20	40.13	41.52	46.84
		$R^2{}_{adj}$	35.32	39.37	41.27	45.70
		5 Total	**35.86**	**40.72**	**40.51**	**45.53**
	10	Distance correlation	40.76	46.71	39.75	45.95
		Pearson correlation	37.34	40.89	40.00	45.44
		$R^2{}_{adj}$	36.71	40.25	39.62	44.30
		10 Total	**38.27**	**42.62**	**39.79**	**45.23**
	25	Distance correlation	38.48	43.67	38.61	42.53
		Pearson correlation	39.75	43.16	42.15	45.06
		$R^2{}_{adj}$	39.37	42.53	41.90	43.92
		25 Total	**39.20**	**43.12**	**40.89**	**43.84**
	50	Distance correlation	39.24	44.81	40.51	47.85
		Pearson correlation	42.53	46.46	40.76	46.58
		$R^2{}_{adj}$	41.90	45.82	40.63	45.44
Probability density function		**50 Total**	**41.22**	**45.70**	**40.63**	**46.62**
	100	Distance correlation	38.35	42.41	41.01	47.22
		Pearson correlation	38.99	42.41	39.62	44.68
		$R^2{}_{adj}$	38.35	41.65	39.37	43.54
		100 Total	**38.57**	**42.15**	**40.00**	**45.15**
	200	Distance correlation	40.38	44.68	40.63	46.84
		Pearson correlation	41.27	43.92	40.25	45.44
		$R^2{}_{adj}$	40.51	43.16	39.87	44.30
		200 Total	**40.72**	**43.92**	**40.25**	**45.53**
	500	Distance correlation	37.09	41.14	38.61	45.44
		Pearson correlation	36.84	39.37	34.56	41.01
		$R^2{}_{adj}$	36.08	38.61	34.18	39.87
		500 Total	**36.67**	**39.70**	**35.78**	**42.11**
	1000	Distance correlation	36.33	42.03	34.05	41.65
		Pearson correlation	34.56	37.59	33.16	39.87
		$R^2{}_{adj}$	34.43	36.84	32.78	38.73
		1000 Total	**35.11**	**38.82**	**33.33**	**40.08**
	2000	Distance correlation	38.10	41.52	33.67	39.37
		Pearson correlation	34.68	35.44	33.29	37.22
		$R^2{}_{adj}$	33.92	34.68	32.91	36.08
		2000 Total	**35.57**	**37.22**	**33.29**	**37.55**
	FD	Distance correlation	48.86	55.32	56.58	62.28

		Pearson correlation	60.63	65.19	60.89	65.32
		R^2_{adj}	59.62	64.30	60.51	64.18
		FD's Total	**54.10**	**56.37**	**61.60**	**59.32**
Optimum number of bins rule	Scott	Distance correlation	52.03	55.70	61.52	64.05
		Pearson correlation	62.53	64.94	65.44	67.47
		R^2_{adj}	62.53	64.18	65.19	66.33
		Scott's Total	**55.37**	**59.03**	**61.60**	**64.05**
	Sturges	Distance correlation	46.20	55.19	52.03	59.37
		Pearson correlation	45.95	53.92	53.92	61.27
		R^2_{adj}	45.95	53.16	53.54	60.13
		Sturges' Total	**45.78**	**46.03**	**54.09**	**53.16**
Discrete empirical distribution	2	Distance correlation	46.08	53.92	58.10	62.66
		Pearson correlation	60.76	67.09	68.10	71.65
		R^2_{adj}	59.37	66.20	67.85	70.51
		2 Total	**55.40**	**62.41**	**64.68**	**68.27**
	5	Distance correlation	49.87	56.96	62.91	68.48
		Pearson correlation	63.04	68.35	70.00	73.92
		R^2_{adj}	62.28	67.59	69.62	72.78
		5 Total	**58.40**	**64.30**	**67.51**	**71.73**
	10	Distance correlation	49.62	57.59	62.15	65.44
		Pearson correlation	48.35	55.70	63.42	66.71
		R^2_{adj}	47.09	54.81	63.16	65.57
		10 Total	**48.35**	**56.03**	**62.91**	**65.91**
	25	Distance correlation	48.10	55.70	57.59	60.38
		Pearson correlation	47.09	52.53	57.09	58.35
		R^2_{adj}	46.71	51.77	56.84	57.22
		25 Total	**47.30**	**53.33**	**57.17**	**58.65**
	50	Distance correlation	42.91	47.85	49.87	53.54
		Pearson correlation	45.82	49.62	51.39	53.80
		R^2_{adj}	46.08	48.99	51.27	52.66
		50 Total	**44.94**	**48.82**	**50.84**	**53.33**
	100	Distance correlation	42.66	47.72	48.48	51.52
		Pearson correlation	50.25	54.56	47.97	50.13
		R^2_{adj}	49.62	53.80	47.72	48.99
		100 Total	**47.51**	**52.03**	**48.06**	**50.21**
	200	Distance correlation	44.94	47.97	47.22	49.87
		Pearson correlation	49.62	51.01	46.08	48.61
		R^2_{adj}	48.73	50.25	45.70	47.47
		200 Total	**47.76**	**49.75**	**46.33**	**48.65**
	500	Distance correlation	38.99	43.29	50.51	52.91

	Pearson correlation	45.32	47.34	50.13	51.52
	$R^2{}_{adj}$	44.43	46.58	49.75	50.38
	500 Total	**42.91**	**45.74**	**50.13**	**51.60**
1000	Distance correlation	40.63	43.67	47.22	50.89
	Pearson correlation	45.32	45.82	46.96	49.62
	$R^2{}_{adj}$	44.43	45.06	46.58	48.48
	1000 Total	**43.46**	**44.85**	**46.92**	**49.66**
2000	Distance correlation	38.99	42.41	45.57	48.61
	Pearson correlation	41.52	43.80	45.32	47.09
	R2adj	40.63	43.04	44.94	45.95
	2000 Total	**40.38**	**43.08**	**45.27**	**47.22**

Table 24 shows the results summarized by calculation method, number of replications, measure of dependence, and impact on the output, where the green color represents the highest value per measure of dependence within calculation method and red represents the lowest values per measure of dependence within calculation method. Regardless of the method used to calculate the MI, the measure of dependence to which the MI is being compared to, or the output that is being investigated, 10 is the number of replications that leads to the worst performance of the MI. This is expected, as with 10 replications there is less information in the simulation generated inputs and, therefore, the reduction in uncertainty of the output provided by the inputs is also lower and, consequently, the ability of the MI to accurately detect the input with the greatest or least impact on the output. With an increase in the number of replications, the performance improves but it also depends on the number of bins, as there is an interaction between number of bins and number of replications as shown in Table 21. When the MI is calculated using fixed number of bins with probability density function, overall, 800 and

1000 are the numbers of replications that lead to the best performance of the MI. When the MI is calculated using fixed number of bins with discrete empirical distribution, 100 and 200 are the numbers of replications that lead to the best performance of the MI. When optimum number of bins rule is used, the number of replications that leads to the best performance of the MI varies between 50 and 200 for the output NIS and is 50 for the output TIS. Although the values are lower when compared to the other methods, it is important to remember that there is an interaction between number of bins and number of replications. When only a lower number of bins (between 2 and 10) is considered for the fixed number of bins method, either with probability density function or discrete empirical distribution, the number of replications that leads to the best performance is also lower. For the first, the number of replications is between 100 and 400, and for the second it is between 20 and 100. In general, the performance of the MI when compared to ${R^2}_{adj}$ or Pearson correlation is more similar than when compared to distance correlation, which is expected as ${R^2}_{adj}$ is a function of R^2 that is equal to the square of the Pearson correlation between the observed y and the predicted values of y. It is worth noting that there are a few differences in terms of the number of replications that leads to the best performance among the different outputs. This is important to highlight because this indicates that depending on the output of interest, the optimum number of replications to be run to reduce the uncertainty on the simulation output can be different.

Table 24. Results from the comparison of the measures of dependence versus the MI summarized by calculation method, number of replications, measure of dependence, and impact on the output considering non-normalized version only.

Calculation method	Number of replications	Measure of dependence	Greatest impact on NIS [%]	Least impact on NIS [%]	Greatest impact on TIS [%]	Least impact on TIS [%]
Probability density function	10	Distance correlation	14.68	17.22	15.32	18.10
		Pearson correlation	13.42	5.70	14.94	11.77
		R^2_{adj}	11.14	6.08	11.77	5.44
		10 Total	**13.08**	**9.66**	**14.01**	**11.77**
	20	Distance correlation	23.04	27.34	23.42	28.99
		Pearson correlation	23.16	17.85	23.92	24.81
		R^2_{adj}	19.37	16.58	23.92	19.75
		20 Total	**21.86**	**20.59**	**23.76**	**24.51**
	50	Distance correlation	35.82	40.25	39.11	43.16
		Pearson correlation	35.70	40.13	40.00	44.18
		R^2_{adj}	35.70	36.33	40.00	44.18
		50 Total	**35.74**	**38.90**	**39.70**	**43.84**
	100	Distance correlation	47.72	51.01	41.77	47.47
		Pearson correlation	42.66	46.58	40.25	47.72
		R^2_{adj}	42.66	44.05	40.25	47.72
		100 Total	**44.35**	**47.22**	**40.76**	**47.64**
	200	Distance correlation	42.66	48.10	40.38	47.97
		Pearson correlation	42.03	47.47	42.66	49.37
		R^2_{adj}	42.03	47.47	42.66	49.37
		200 Total	**42.24**	**47.68**	**41.90**	**48.90**
	400	Distance correlation	42.03	47.47	41.77	48.73
		Pearson correlation	41.77	47.47	41.77	48.48
		R^2_{adj}	41.77	47.47	41.77	48.48
		400 Total	**41.86**	**47.47**	**41.77**	**48.57**
	600	Distance correlation	41.27	47.85	43.04	50.25
		Pearson correlation	42.53	48.35	43.42	50.25
		R^2_{adj}	42.53	48.35	43.42	50.25
		600 Total	**42.11**	**48.19**	**43.29**	**50.25**
	800	Distance correlation	45.44	51.65	44.56	52.03
		Pearson correlation	45.70	51.52	44.56	52.53
		R^2_{adj}	45.70	51.52	44.56	52.53
		800 Total	**45.61**	**51.56**	**44.56**	**52.36**
	1000	Distance correlation	43.92	49.75	45.32	52.41
		Pearson correlation	45.70	51.27	45.70	53.16

		$R^2{}_{adj}$	45.70	51.27	45.70	53.16
		1000 Total	**45.11**	**50.76**	**45.57**	**52.91**
	1500	Distance correlation	41.65	46.84	42.78	50.51
		Pearson correlation	42.15	47.85	42.28	49.37
		$R^2{}_{adj}$	42.15	47.85	42.28	49.37
		1500 Total	**41.98**	**47.51**	**42.45**	**49.75**
Optimum number of bins rule	10	Distance correlation	33.33	41.77	37.97	41.77
		Pearson correlation	40.51	33.76	42.62	37.55
		$R^2{}_{adj}$	37.13	33.33	39.24	31.22
		10 Total	**36.68**	**36.99**	**36.29**	**39.94**
	20	Distance correlation	52.74	63.71	56.54	68.78
		Pearson correlation	57.38	61.60	58.65	67.51
		$R^2{}_{adj}$	57.38	60.34	58.65	62.45
		20 Total	**53.00**	**55.84**	**61.88**	**57.95**
	50	Distance correlation	56.96	64.14	65.40	70.46
		Pearson correlation	62.45	68.78	69.20	74.68
		$R^2{}_{adj}$	62.45	64.98	69.20	74.68
		50 Total	**55.95**	**60.62**	**65.96**	**67.93**
	100	Distance correlation	56.12	60.34	59.49	64.56
		Pearson correlation	63.29	69.62	63.29	70.04
		$R^2{}_{adj}$	63.29	67.09	63.29	70.04
		100 Total	**58.70**	**60.90**	**65.68**	**62.03**
	200	Distance correlation	48.10	54.01	56.96	62.03
		Pearson correlation	65.40	73.00	64.56	71.73
		$R^2{}_{adj}$	65.40	73.00	64.56	71.73
		200 Total	**56.83**	**59.63**	**66.67**	**62.03**
	400	Distance correlation	43.46	48.52	59.07	64.14
		Pearson correlation	62.03	68.78	65.40	70.89
		$R^2{}_{adj}$	62.03	68.78	65.40	70.89
		400 Total	**51.03**	**55.84**	**62.03**	**63.29**
	600	Distance correlation	45.99	53.59	56.12	62.45
		Pearson correlation	57.38	65.40	63.29	69.20
		$R^2{}_{adj}$	57.38	65.40	63.29	69.20
		600 Total	**52.31**	**53.59**	**61.46**	**60.90**
	800	Distance correlation	51.48	57.38	58.65	63.29
		Pearson correlation	54.85	61.18	60.34	65.40
		$R^2{}_{adj}$	54.85	61.18	60.34	65.40
		800 Total	**51.72**	**53.73**	**59.92**	**59.77**
	1000	Distance correlation	50.21	54.01	59.07	61.60
		Pearson correlation	48.95	55.27	58.65	61.60

		$R^2{}_{adj}$	48.95	55.27	58.65	61.60
		1000 Total	**49.75**	**49.37**	**54.85**	**58.79**
	1500	Distance correlation	51.90	56.54	57.81	59.92
		Pearson correlation	51.48	56.12	54.85	58.23
		$R^2{}_{adj}$	51.48	56.12	54.85	58.23
		1500 Total	**51.52**	**51.62**	**56.26**	**55.84**
Discrete empirical distribution	10	Distance correlation	16.33	20.89	20.38	22.91
		Pearson correlation	16.08	9.11	20.89	16.71
		$R^2{}_{adj}$	13.29	8.99	17.85	10.38
		10 Total	**15.23**	**13.00**	**19.70**	**16.67**
	20	Distance correlation	37.85	41.52	40.38	42.91
		Pearson correlation	39.24	33.29	41.01	38.99
		$R^2{}_{adj}$	34.30	32.03	41.01	33.92
		20 Total	**37.13**	**35.61**	**40.80**	**38.61**
	50	Distance correlation	50.00	53.42	54.81	58.35
		Pearson correlation	50.76	53.80	56.08	59.62
		$R^2{}_{adj}$	50.76	50.00	56.08	59.62
		50 Total	**50.51**	**52.41**	**55.65**	**59.20**
	100	Distance correlation	54.43	58.10	62.78	64.81
		Pearson correlation	60.76	65.32	64.05	67.09
		$R^2{}_{adj}$	60.76	62.78	64.05	67.09
		100 Total	**58.65**	**62.07**	**63.63**	**66.33**
	200	Distance correlation	50.00	54.81	60.76	63.92
		Pearson correlation	58.73	63.16	64.18	67.47
		$R^2{}_{adj}$	58.73	63.16	64.18	67.47
		200 Total	**55.82**	**60.38**	**63.04**	**66.29**
	400	Distance correlation	46.96	52.53	56.46	59.37
		Pearson correlation	54.56	60.00	62.03	65.57
		$R^2{}_{adj}$	54.56	60.00	62.03	65.57
		400 Total	**52.03**	**57.51**	**60.17**	**63.50**
	600	Distance correlation	43.80	51.01	57.72	61.77
		Pearson correlation	51.90	59.62	58.35	62.91
		$R^2{}_{adj}$	51.90	59.62	58.35	62.91
		600 Total	**49.20**	**56.75**	**58.14**	**62.53**
	800	Distance correlation	44.81	54.30	57.97	64.30
		Pearson correlation	53.67	64.56	58.23	64.56
		$R^2{}_{adj}$	53.67	64.56	58.23	64.56
		800 Total	**50.72**	**61.14**	**58.14**	**64.47**
	1000	Distance correlation	48.35	54.18	59.37	63.54
		Pearson correlation	56.46	64.18	61.27	64.56

	$R^2{}_{adj}$	56.46	64.18	61.27	64.56
	1000 Total	**53.76**	**60.84**	**60.63**	**64.22**
1500	Distance correlation	50.25	56.33	58.99	62.41
	Pearson correlation	54.94	62.78	60.38	63.92
	$R^2{}_{adj}$	54.94	62.78	60.38	63.92
	1500 Total	**53.38**	**60.63**	**59.92**	**63.42**

Table 25 shows the results summarized by calculation method, measure of dependence, and impact on the output, where the green color represents the highest value per impact on the output within calculation method and red represents the lowest values per impact on the output within calculation method. As shown in Table 25, overall, regardless of the method used to calculate the MI and the measure of dependence to which the MI is being compared to, the MI appears to have worst performance in detecting the input that has the greatest impact on the NIS and best performance in detecting the input that has the least impact on the TIS. In general, the simulation modeler should be more interested in knowing the input that has the greatest impact on the output than the input that has the least impact on the output, as this will allow the simulation modeler to better plan the limited resources for data collection and for running the experiments in order to reduce uncertainty in the simulation model. On the other hand, knowing the input that has the least impact on the output allows the simulation modeler to eliminate inputs that are not as valuable in case of limited resources.

Among the outputs considered, MI showed better performance in detecting the input with either the least or the greatest impact on the TIS than on the NIS. Considering

only fixed number of bins, this difference is more evident when using discrete empirical distribution than probability density function. There is a possible explanation for the difference. NIS is a simulation output that is an average over time and TIS is an average over a number of entities that represent the customers in the system. When using the histogram-based method, whether using the discrete empirical distribution or the probability density function, information is lost by "binning". Therefore, the discontinuity of the histogram may explain the better performance for the output that has a discrete behavior. The difference could also have different explanations: more replications are needed in order to improve the performance for the NIS or the inputs in the system being studied have a stronger relationship with TIS than with NIS, which makes it easier to detect their relationship in either the low or high level.

Table 25. Results from the comparison of the measures of dependence versus the MI summarized by calculation method, measure of dependence, and impact on the output considering non-normalized version only.

Impact on the output	Calculation method	Distance correlation [%]	Pearson correlation [%]	$R^2{}_{adj}$ [%]	Total [%]
Greatest impact on NIS	Discrete empirical distribution	44.28	49.71	48.94	47.64
	Probability density function	37.82	37.48	36.87	37.39
	Optimum number of bins rule	49.03	56.37	56.03	53.81
Greatest impact on NIS Total		**41.94**	**42.09**	**45.26**	**44.62**
Least impact on NIS	Discrete empirical distribution	49.71	53.58	52.81	52.03
	Probability density function	42.75	40.42	39.70	40.95
	Optimum number of bins rule	56.71	60.08	59.75	58.85
Least impact on NIS Total		**47.02**	**46.84**	**48.10**	**47.78**
Greatest impact on TIS	Discrete empirical distribution	52.96	54.65	54.34	53.98
	Probability density function	37.75	37.95	37.63	37.78
	Optimum number of bins rule	55.40	61.35	60.55	59.10

Greatest impact on TIS Total		**47.43**	**47.42**	**48.87**	**48.12**
Least impact on TIS	Discrete empirical distribution	56.43	57.14	56.00	56.52
	Probability density function	43.96	43.16	42.03	43.05
	Optimum number of bins rule	61.90	64.68	63.54	63.38
Least impact on TIS Total		**52.13**	**51.72**	**52.05**	**50.91**

Finally, χ^2 test at an α-level of 0.05 was performed to investigate whether the performance of the MI was statistically significantly different based on the method used to calculate the MI, the dependence measure to which the MI was being compared to, and the output that was being investigated, as shown in Table 26, Table 27, and Table 28, respectively.

Table 26 shows the χ^2 test results whether the performance of the MI is statistically significantly different based on the calculation method. For this test, only the non-normalized version of the MI is considered. The null hypothesis is that the MI performance based on the two calculation methods are not different and the alternative hypothesis is that the MI performance based on the two calculation methods are different. When the MI is calculated using fixed bins with discrete empirical distribution or optimum number of bins rule, the measure shows statistically significantly better performance than when it is calculated using fixed bins with probability density function, regardless of the impact on the output. The optimum number of bins rule is the method that led to the overall best performance of the MI measure regardless of the impact on the output and the measure of dependence to which the MI was being compared to. This unexpected better performance of the optimum number of bins rule method, different than what was observed for the

entropy measures, can be possibly explained due to the fact that the MI performance is better for lower values of bins. In this work the optimum number of bins rule method contains only lower number of bins, consequently its results are better.

Table 26. χ^2 test results whether the performance of the MI is statistically significantly different based on the calculation method.

Measure of dependence	Impact on output	Proportion difference	P-value
Distance correlation	Greatest impact on NIS	0.0646 (DED[1]-PDF[2])	**<0.0001**
	Least impact on NIS	0.0696 (DED-PDF)	**<0.0001**
	Greatest impact on TIS	0.1522 (DED-PDF)	**<0.0001**
	Least impact on TIS	0.1247 (DED-PDF)	**<0.0001**
Pearson correlation	Greatest impact on NIS	0.1223 (DED-PDF)	**<0.0001**
	Least impact on NIS	0.1316 (DED-PDF)	**<0.0001**
	Greatest impact on TIS	0.1670 (DED-PDF)	**<0.0001**
	Least impact on TIS	0.1397 (DED-PDF)	**<0.0001**
$R^2{}_{adj}$	Greatest impact on NIS	0.1206 (DED-PDF)	**<0.0001**
	Least impact on NIS	0.1311 (DED-PDF)	**<0.0001**
	Greatest impact on TIS	0.1671 (DED-PDF)	**<0.0001**
	Least impact on TIS	0.1397 (DED-PDF)	**<0.0001**
Distance correlation	Greatest impact on NIS	0.1121 (OPT[3]-PDF)	**<0.0001**
	Least impact on NIS	0.1265 (OPT-PDF)	**<0.0001**
	Greatest impact on TIS	0.1896 (OPT-PDF)	**<0.0001**
	Least impact on TIS	0.1794 (OPT-PDF)	**<0.0001**
Pearson correlation	Greatest impact on NIS	0.1889 (OPT-PDF)	**<0.0001**
	Least impact on NIS	0.2093 (OPT-PDF)	**<0.0001**
	Greatest impact on TIS	0.2214 (OPT-PDF)	**<0.0001**
	Least impact on TIS	0.2152 (OPT-PDF)	**<0.0001**
$R^2{}_{adj}$	Greatest impact on NIS	0.1916 (OPT-PDF)	**<0.0001**
	Least impact on NIS	0.2085 (OPT-PDF)	**<0.0001**
	Greatest impact on TIS	0.2211 (OPT-PDF)	**<0.0001**
	Least impact on TIS	0.2152 (OPT-PDF)	**<0.0001**
Distance correlation	Greatest impact on NIS	-0.0475 (DED-OPT)	**<0.0001**
	Least impact on NIS	-0.0569 (DED-OPT)	**<0.0001**
	Greatest impact on TIS	-0.0375 (DED-OPT)	**0.0013**
	Least impact on TIS	-0.0547 (DED-OPT)	**<0.0001**

Pearson correlation	Greatest impact on NIS	-0.0666 (DED-OPT)	**<0.0001**
	Least impact on NIS	-0.0777 (DED-OPT)	**<0.0001**
	Greatest impact on TIS	-0.0544 (DED-OPT)	**<0.0001**
	Least impact on TIS	-0.0754 (DED-OPT)	**<0.0001**
$R^2{}_{adj}$	Greatest impact on NIS	-0.0710 (DED-OPT)	**<0.0001**
	Least impact on NIS	-0.0774 (DED-OPT)	**<0.0001**
	Greatest impact on TIS	-0.0541 (DED-OPT)	**<0.0001**
	Least impact on TIS	-0.0754 (DED-OPT)	**<0.0001**

[1]DED is the fixed bins with discrete empirical distribution method

[2]PDF is the fixed bins with probability density function method

[3]OPT is the optimum number of bins rule with probability density function method

Table 27 shows the χ^2 test results whether the performance of the MI is statistically significantly different based on the measure of dependence. The null hypothesis is that there is no difference in the MI performance among the two measures of dependence and the alternative hypothesis is that the two measures of dependence lead to differences in MI performance. Regardless of the calculation method, there is no evidence of difference in the results of the MI performance when MI is being compared to Pearson correlation versus when MI is compared to $R^2{}_{adj}$. When optimum number of bins rule is the method used to calculate the MI, the χ^2 test shows that the MI is statistically significantly better when compared to Pearson correlation and $R^2{}_{adj}$ than when compared to distance correlation. When fixed number of bins with probability density function is the method used to calculate the MI, the χ^2 test only shows statistically significantly difference in the MI performance to detect the input with the least impact on the NIS or the TIS. In these cases, MI showed better performance when compared to distance correlation than when

compared to Pearson correlation or $R^2{}_{adj}$. When fixed number of bins with discrete empirical distribution is the method used to calculate the MI, the χ^2 test results are slightly different based on the normalization used. If no normalization is considered the χ^2 test shows statistically significantly difference in the MI performance to detect the input with the least or the greatest impact on the NIS. When normalization is considered (except NMI_{theor}), the statistically significantly difference in performance only occurs for detecting the input with the greatest impact on the NIS. In these cases, MI showed better performance when compared to Pearson correlation or $R^2{}_{adj}$ than when compared to distance correlation, which is the opposite to what was observed when using probability density function. A possible explanation for the difference in the χ^2 test results when using the probability density function and the discrete empirical distribution is that when using probability density function the possible non-linear relation of the inputs and outputs has been better captured than when using discrete empirical distribution and, hence, this first calculation method has a better performance when compared to distance correlation than when compared to Pearson or $R^2{}_{adj}$.

Table 27. χ^2 test results whether the performance of MI is statistically significantly different based on the measure of dependence.

Method of calculation	Normalization method	Impact on output	Proportion difference	P-value
Probability density function	Non-normalization	Greatest impact on NIS	0.0034 (distance correlation - Pearson[1])	0.6575
		Least impact on NIS	0.0233 (distance correlation - Pearson)	**0.0030**

		Greatest impact on TIS	-0.0020 (distance correlation - Pearson)	0.7930
		Least impact on TIS	0.0080 (distance correlation - Pearson)	0.3121
		Greatest impact on NIS	0.0095 (distance correlation - R^2_{adj})	0.2174
		Least impact on NIS	0.0305 (distance correlation - R^2_{adj})	**<0.0001**
		Greatest impact on TIS	0.0011 (distance correlation - R^2_{adj})	0.8825
		Least impact on TIS	0.0194 (distance correlation - R^2_{adj})	**0.0139**
		Greatest impact on NIS	0.0061 (Pearson - R^2_{adj})	0.4294
		Least impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3547
		Greatest impact on TIS	0.0032 (Pearson - R^2_{adj})	0.6817
		Least impact on TIS	0.0114 (Pearson - R^2_{adj})	0.1476
Optimum number of bins rule	Non-normalization	Greatest impact on NIS	-0.0734 (distance correlation - Pearson)	**<0.0001**
		Least impact on NIS	-0.0595 (distance correlation - Pearson)	**<0.0001**
		Greatest impact on TIS	-0.0338 (distance correlation - Pearson)	**0.0184**
		Least impact on TIS	-0.0278 (distance correlation - Pearson)	**0.0467**
		Greatest impact on NIS	-0.0700 (distance correlation - R^2_{adj})	**<0.0001**
		Least impact on NIS	-0.0515 (distance correlation - R^2_{adj})	**0.0003**
		Greatest impact on TIS	-0.0304 (distance correlation - R^2_{adj})	**0.0340**
		Least impact on TIS	-0.0165 (distance correlation - R^2_{adj})	0.2416
		Greatest impact on NIS	0.0034 (Pearson - R^2_{adj})	0.8149
		Least impact on NIS	0.0080 (Pearson - R^2_{adj})	0.5718
		Greatest impact on TIS	0.0034 (Pearson - R^2_{adj})	0.8127
		Least impact on TIS	0.0114 (Pearson - R^2_{adj})	0.4138
Discrete empirical distribution	Non-normalization	Greatest impact on NIS	-0.0543 (distance correlation - Pearson)	**<0.0001**
		Least impact on NIS	-0.0387 (distance correlation - Pearson)	**<0.0001**
		Greatest impact on TIS	-0.0168 (distance correlation - Pearson)	**0.0338**
		Least impact on TIS	-0.0071 (distance correlation - Pearson)	0.3685
	Arith	Greatest impact on NIS	-0.0322 (distance correlation - Pearson)	**<0.0001**

	Least impact on NIS	-0.0149 (distance correlation - Pearson)	0.0604
	Greatest impact on TIS	-0.0139 (distance correlation - Pearson)	0.0794
	Least impact on TIS	-0.0058 (distance correlation - Pearson)	0.4605
Joint	Greatest impact on NIS	-0.0323 (distance correlation - Pearson)	**<0.0001**
	Least impact on NIS	-0.0149 (distance correlation - Pearson)	0.0604
	Greatest impact on TIS	-0.0141 (distance correlation - Pearson)	0.0767
	Least impact on TIS	-0.0058 (distance correlation - Pearson)	0.4605
Geom	Greatest impact on NIS	-0.0322 (distance correlation - Pearson)	**<0.0001**
	Least impact on NIS	-0.0149 (distance correlation - Pearson)	0.0604
	Greatest impact on TIS	-0.0135 (distance correlation - Pearson)	0.0880
	Least impact on TIS	-0.0053 (distance correlation - Pearson)	0.5004
Theor	Greatest impact on NIS	-0.0544 (distance correlation - Pearson)	**<0.0001**
	Least impact on NIS	-0.0389 (distance correlation - Pearson)	**<0.0001**
	Greatest impact on TIS	-0.0170 (distance correlation - Pearson)	**0.0325**
	Least impact on TIS	-0.0072 (distance correlation - Pearson)	0.3600
Non-normalization	Greatest impact on NIS	-0.0466 (distance correlation - R^2_{adj})	**<0.0001**
	Least impact on NIS	-0.0310 (distance correlation - R^2_{adj})	**<0.0001**
	Greatest impact on TIS	-0.0138 (distance correlation - R^2_{adj})	0.0820
	Least impact on TIS	0.0043 (distance correlation - R^2_{adj})	0.5856
Arith	Greatest impact on NIS	-0.0249 (distance correlation - R^2_{adj})	**0.0017**
	Least impact on NIS	-0.0072 (distance correlation - R^2_{adj})	0.3644
	Greatest impact on TIS	-0.0110 (distance correlation - R^2_{adj})	0.1654
	Least impact on TIS	0.0056 (distance correlation - R^2_{adj})	0.4808
Joint	Greatest impact on NIS	-0.0249 (distance correlation - R^2_{adj})	**0.0017**

	Least impact on NIS	-0.0072 (distance correlation - R^2_{adj})	0.3644
	Greatest impact on TIS	-0.0111 (distance correlation - R^2_{adj})	0.1606
	Least impact on TIS	0.0056 (distance correlation - R^2_{adj})	0.4808
Geom	Greatest impact on NIS	-0.0249 (distance correlation - R^2_{adj})	**0.0017**
	Least impact on NIS	-0.0072 (distance correlation - R^2_{adj})	0.3644
	Greatest impact on TIS	-0.0106 (distance correlation - R^2_{adj})	0.1805
	Least impact on TIS	0.0061 (distance correlation - R^2_{adj})	0.4419
Theor	Greatest impact on NIS	-0.0467 (distance correlation - R^2_{adj})	**<0.0001**
	Least impact on NIS	-0.0311 (distance correlation - R^2_{adj})	**<0.0001**
	Greatest impact on TIS	-0.0139 (distance correlation - R^2_{adj})	0.0793
	Least impact on TIS	0.0042 (distance correlation - R^2_{adj})	0.5967
Non-normalization	Greatest impact on NIS	0.0077 (Pearson - R^2_{adj})	0.3317
	Least impact on NIS	0.0077 (Pearson - R^2_{adj})	0.3308
	Greatest impact on TIS	0.0030 (Pearson - R^2_{adj})	0.7014
	Least impact on TIS	0.0114 (Pearson - R^2_{adj})	0.1486
Arith	Greatest impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3638
	Least impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3315
	Greatest impact on TIS	0.0029 (Pearson - R^2_{adj})	0.7136
	Least impact on TIS	0.0114 (Pearson - R^2_{adj})	0.1490
Joint	Greatest impact on NIS	0.0073 (Pearson - R^2_{adj})	0.3555
	Least impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3315
	Greatest impact on TIS	0.0029 (Pearson - R^2_{adj})	0.7135
	Least impact on TIS	0.0114 (Pearson - R^2_{adj})	0.1490
Geom	Greatest impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3638
	Least impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3315
	Greatest impact on TIS	0.0029 (Pearson - R^2_{adj})	0.7136
	Least impact on TIS	0.0114 (Pearson - R^2_{adj})	0.1491
Theor	Greatest impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3317
	Least impact on NIS	0.0072 (Pearson - R^2_{adj})	0.3308
	Greatest impact on TIS	0.0030 (Pearson - R^2_{adj})	0.7014
	Least impact on TIS	0.0114 (Pearson - R^2_{adj})	0.1486

[1]Pearson is the Pearson correlation measure of dependence.

Table 28 shows the χ^2 test results whether the performance of the MI is statistically significantly different based on the output being investigated. The null hypothesis is that there is no difference in the MI performance among the outputs being investigated and the alternative hypothesis is that the outputs being investigated lead to differences in MI performance.

As shown in Table 28, with one exception, regardless of the method used to calculate the MI, the normalization method, and the impact on the output, there is statistically significantly difference in the MI performance based on the output being investigated. The exception occurs when the MI is calculated using fixed number of bins with probability density function to detect the input with the greatest impact on the output. The MI has statistically significantly better performance in detecting the impact on the TIS than on the NIS. This has also been previously observed in the results shown in Table 25.

Table 28. χ^2 test results whether the performance of MI is statistically significantly different based on the output being investigated.

Method of calculation	Normalization method	Impact on output	Proportion difference	P-value
Probability density function	Non-normalization	Greatest impact	-0.0038 (NIS-TIS)	0.3881
		Least impact	-0.0210 (NIS-TIS)	**<0.0001**
Optimum number of bins rule	Non-normalization	Greatest impact	-0.0504 (NIS-TIS)	**<0.0001**
		Least impact	-0.0428 (NIS-TIS)	**<0.0001**

Discrete empirical distribution	Non-normalization	Greatest impact	-0.0634 (NIS-TIS)	**<0.0001**
		Least impact	-0.0449 (NIS-TIS)	**<0.0001**
	Arith	Greatest impact	-0.0702 (NIS-TIS)	**<0.0001**
		Least impact	-0.0495 (NIS-TIS)	**<0.0001**
	Joint	Greatest impact	-0.0703 (NIS-TIS)	**<0.0001**
		Least impact	-0.0495 (NIS-TIS)	**<0.0001**
	Geom	Greatest impact	-0.0694 (NIS-TIS)	**<0.0001**
		Least impact	-0.0491 (NIS-TIS)	**<0.0001**
	Theor	Greatest impact	-0.0637 (NIS-TIS)	**<0.0001**
		Least impact	-0.0452 (NIS-TIS)	**<0.0001**

Figure 29 shows the results of MI performance compared to the different measures of dependence, per different number of bins, per different number of replications, per method of calculation, and per output. Figure 29 shows some of the observations made in this section in a more visual way for easier comparison. For instance, with respect to the comparison between the MI performance among the discrete empirical distribution and the probability density function or among the different number of replications.

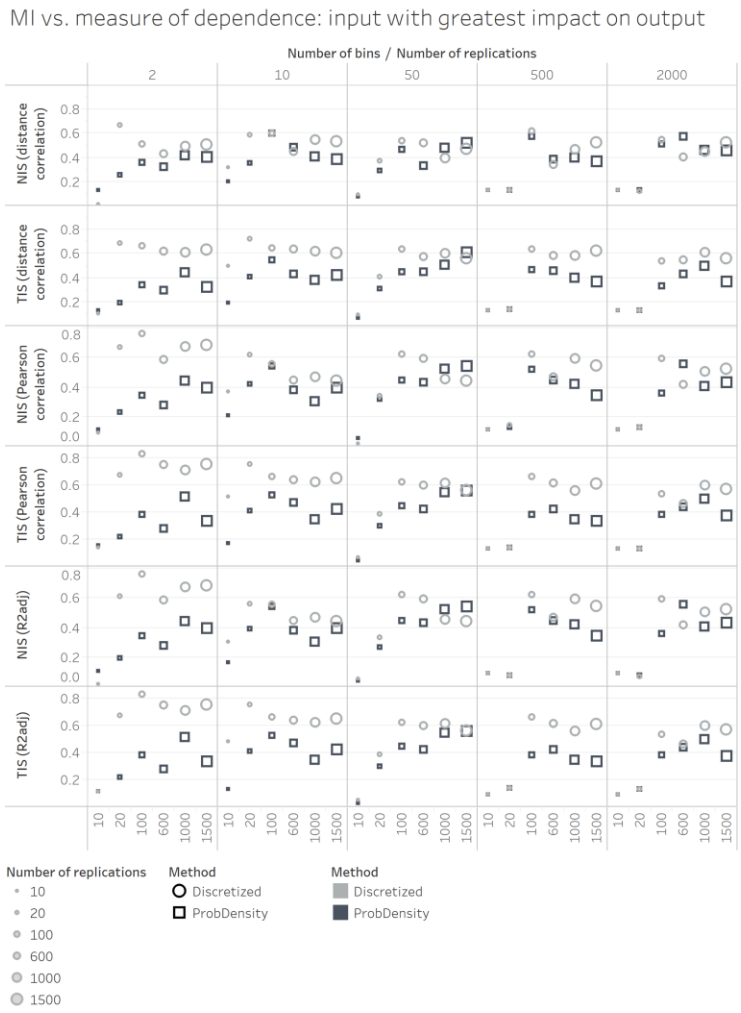

Figure 29. Results for MI vs. measure of dependence for detecting the input with the greatest impact on the output, per number of bins, number of replications, method of calculation, measure of dependence, and output.

2.5. Concluding remarks

Although defining uncertainty quantification is simple, developing a systematic method is difficult and hard to validate. In this work, through a total of 1,130 experiments, Shannon's entropy and MI calculated using the histogram-based method were investigated

as potential measures of uncertainty quantification in simulation model. The first contribution of this section of the dissertation was to discuss the challenges found while applying entropy measures for continuous variables and to identify a few issues of interpretability faced when using the method proposed by Jaynes (1957) with $m(x) = sup\left[\hat{f}(x)\right]$. Based on the issues, it was showed that when using fixed number of bins, changing the data normalization does not change the placement of the data into bins and an alternative to calculate the entropy and the MI measures for continuous variables was proposed. This alternative involved normalizing the data in a way not only to avoid the differences of spread in the inputs and outputs, but also to guarantee that $0 \leq \hat{f}_i(x) < 1, \forall i$ and, hence, to avoid the issues identified when $m(x) = sup\left[\hat{f}(x)\right]$. However, it is important to note that with this alternative, there is still the issue that the entropy is not maximum when the $\hat{f}_i(x)$ is equiprobable, as $\sum_i \hat{f}_i(x)$ is not necessarily equal to 1 as it is in when the inputs and outputs are discrete.

In section 2.4.2, the impact of different binwidths and different normalization methods on the entropy and MI measures was discussed. An important contribution of this section is that it was showed that when entropy and MI are calculated using histogram-based method with probability density function, the measures tend to decrease with the increase in the number of bins; while when the measures are calculated using histogram-based method with discrete empirical distribution, they tend to increase with the increase in the number of bins. This is important because while the latter is mentioned in the literature, the first was not found to be mentioned in the information theory literature even

after a comprehensive literature review. Due to the impact of the number of bins on the entropy measures, the literature recommends normalizing the entropy and mutual information. When the entropy and MI measures calculated using discrete empirical distribution are normalized by the theoretical maximum, the measures behave in a similar way to the measures calculated with probability density function. Nevertheless, if the MI is normalized by the real maximum, the measure still increases with the increase in the number of bins. For this reason and because the NMI_{theor} measures have similar results of performance to the non-normalized MI version when compared to other measures of dependence, the normalization using the theoretical maximum (NMI_{theor}) is the final recommendation on the basis of this work. Another recommendation is that normalization should be used only for the entropy measures calculated using the discrete empirical distribution, as the entropy measures calculated using the probability density function do not increase with the increase in the number of bins. Normalization when using probability density function is recommended for comparison of MI measures calculated using different number of bins rather than to eliminate the effect of the number of bins.

For the case of the entropy and MI measures calculated using probability density function or the theoretical normalized version of the discrete empirical distribution, the number of bins could be interpreted as the level of accuracy one wants to obtain. This interpretation would explain why entropy and MI decrease with the increase of the number of bins: when someone cares about a greater level of accuracy, the same number of replications should be able to provide less information or less reduction in uncertainty.

Another important point is that the first few replications should bring more information about the system than any subsequent ones. This was observed when the entropy measures were calculated either using the probability density function of the theoretical normalizaed version of the discrete empirical distribution.

From section 2.4.2, it was also possible to observe that when the entropy and MI measures were calculated using either the probability density function or the discrete empirical distribution after normalization, low number of bins (i.e. number of bins between 2 and 10) appeared to be inadequate as they led to results not consistent to results from higher number of bins. The results of the entropy and MI measures calculated using the optimum number of bins also leads to this conclusion, because most of these results were not consistent and they were calculated with a low number of bins. Therefore, a recommendation would be to use a number of bins of at least 25 to calculate the entropy and MI measures.

In section 2.4.3, the impact of different traffic intensities, different seeds, different parameter values, and different systems on the entropy and MI measures was investigated. The main observations from this section were that regardless of the method chosen and the number of bins used: (i) the entropy measure was able to correctly identify that X_1 has the same information/uncertainty among the different traffic intensity experiments; and (ii) the entropy measures indicate differences in information/uncertainty based on different seeds, different traffic intensities, and different parameter values. Although a consistency was not observed on these latter differences, it is important to highlight that the focus was

not to investigate the relationship among the entropy measures and the different seeds, traffic intensities, and parameter values. An interaction among these factors and the model type ($M/M/1$ and $M/G/1$), for instance, could exist and it was not investigated in this work. Another important observation is that when the normalized entropy and NMI_{theor} were calculated using discrete empirical distribution, the measures were able to appropriately point the null uncertainty in $\widehat{Y_1}$ in the CONWIP system and in X_3 when the latter is deterministic, as well as their zero impact in the inputs and outputs, respectively. When the measures were calculated using probability density function, the null uncertainty and the zero impact were only pointed for larger number of bins (when number of bins was greater than or equal to 1,000). Although for lower number of bins, the entropy and MI measures were not able to capture zero uncertainty, they were still able to capture some deterministic behavior, as they were constant regardless of the number of replications used. Therefore, through the results of this section, one can see the potential of the entropy and MI as measures of uncertainty quantification in simulation models, as the measures were able to capture different and important characteristics in the simulation model. However, it is important to highlight that when the entropy and MI measures was calculated using the probability density function, they were only able to fully capture the deterministic behavior for number of bins greater than or equal to 1,000, which could either indicate an issue with the measure or just the fact that the number of bins must be well chosen for this method to work.

Another important conclusion from section 2.4.3 is that entropy and MI may be an alternative method to investigate the quality of a group of seeds in simulation models. As the results indicated, while the entropy and MI of the original group of seeds and group "seed 2" were similar among themselves, they were slightly different from group "seed 3", which could indicate an issue with this latter group of seeds.

In section 2.4.4, the results of the entropy measures were compared to SAE, SSE, MAE, and MSE and the results of the MI measures were compared to distance correlation, Pearson correlation, and $R^2{}_{adj}$ to identify whether the measures agreed with other methods of the literature.

With respect to the entropy measures, the following observation are worth highlighting: (i) when using fixed number of bins with probability density function, overall, the agreement of the entropy measures with SAE, SSE, MAE, and MSE increased with the increase in the number of bins and 1,000 was the number of bins that led to the best results; and, (ii) when using fixed number of bins with discrete empirical distribution, the percentage of the entropy measures that agreed with SAE and SSE increased with the increase in the number of bins, but was constant with MAE and MSE, and, in general, the number of bins between 1,000 and 2,000 led to the best results. Because comparing the measures with SAE and SSE may have an issue of bias, the comparison to MAE and MSE is preferred.

As the χ^2 test showed, for the number of bins that led to the best performance, the probability density function method was statistically significantly better than the discrete

empirical distribution method. Other important conclusions from the entropy measures are that the optimum number of bins rules did not lead to good results compared to the other error measures (e.g., compared to MAE, the best results from this method ranged from 2.6% to 50%, while for the fixed number of bins with probability density function they ranged from 57.9% to 81.1%). Another important observation is that the normalized version of the entropy measure calculated using the histogram-based method with fixed number of bins and discrete empirical distribution led to results identical to the non-normalized version.

With respect to the MI measures, a relevant conclusion made from the logistic regression model was that, overall, number of replications and interaction between number of replications and number of bins were the most important factors in the performance of the MI when compared to the other measures of dependence, regardless of the method used to calculate the MI.

Based on the comparison of the MI to the other measures of dependences, a few important observations could be made: (i) when using probability density function, low or high-range number of bins had the worst performance and mid-range values had the best performance, while using discrete empirical distribution, 2 had the worst performance and 5 had the best performance; (ii) when using optimum number of bins rules, Sturges' led to the worst performance, as expected, and Scott's led to the best performance; and, (iii) when using probability density function, 800 and 1,000 were the number of replications that led to the best performance, while using discrete empirical distribution, 100 and 200

replications led to the best performance, and for optimum number of bins it was between 50 and 200. Regarding the first observation about the number of bins, this result is different from the entropy measure, where in general 1,000 was the number of bins that led to the best results. However, it is important to note that there is an interaction between number of bins and number of replications, which could explain the differences. Other explanations for the differences are: (1) the fact that entropy and MI measures measure different concepts (i.e., uncertainty of the inputs and outputs vs. the impact of the input on the output and vice-versa), and, (2) the fact that the entropy and MI measures were compared to other measures of uncertainty and dependence that may not have been the most adequate comparison. For instance, distance correlation was the only measure capable of measuring non-linear dependence and according to the literature, MI is also capable of measuring non-linear dependence. Therefore, comparing MI with other measures that can only measure linear dependence may not be the most appropriate alternative.

Regarding the differences in the number of replications, they can be possibly explained by the fact that the optimum number of bins comprises smaller number of bins experiments. In this case, if one thinks about number of bins as level of accuracy, one should be able to obtain the optimum amount of information about the system more quickly. Another possible explanation is that when using the discrete empirical distribution, one is approximating the continuous data with a discrete method. Therefore, one should also reach the maxima more quickly than when using the probability density

function, as it is just an approximation. It is also important highlighting that there are a few differences in terms of the number of replications that leads to the best performance among the different outputs. This is important to note because if one is using the MI measures as an alternate method to determine the number of replications to run to reduce the uncertainty on the simulation output, the optimum number of replications to be run could be different based on the output of interest.

Different than what was observed for the entropy measures, the MI showed better performance when calculated using optimum number of bins rule, followed by the discrete empirical distribution and the probability density function with fixed number of bins, respectively. This result is again contrary to what was observed in the entropy measures, which is another indication that either the comparison of the MI with the measures of the dependence may not be the best alternative or the comparison of the entropy with the measures of error may not be adequate.

Another valuable insight from section 2.4.4 is that, as expected, the results of the MI performance when MI was compared to Pearson were not statistically significantly different than when MI was compared to ${R^2}_{adj}$. This is expected as ${R^2}_{adj}$ is a function of R^2 that is equal to the square of the Pearson correlation between the observed y and the predicted values of y. Overall, when the MI measures were calculated using the probability density function, they showed better performance when compared to the distance correlation measure. When the MI measures were calculated using optimum number of bins rule or discrete empirical distribution with normalization, the MI

performance was in general statistically significantly better when compared to Pearson correlation and $R^2{}_{adj}$ than when compared to distance correlation. This may indicate that when the MI measures are calculated using the discrete empirical distribution or low number of bins, the non-linear relation of the inputs and outputs is not being captured.

As for the main limitations of this work, there are: (i) histogram-based method was the only method used to calculate the entropy and MI measures; and (ii) although the method was compared with other well-known measures of the scientific community, the method was not validated theoretically. Nevertheless, as showed in this work, despite the challenges encoutered in the application, entropy and mutual information measures present good and promising results in measuring uncertainty in simulation model.

A recommendation for future research is to investigate how other methods, such as the kernel-based method and the k-nearest neighbors method would affect the results.

A question one could ask is: “Given the challenges discussed and the lack of theoretical validity, why would someone choose to use the proposed entropy method for uncertainty quantification instead of simply using the well-known confidence intervals based on the standard error of the estimator?” First, confidence intervals do not quantify uncertainty. That is, confidence intervals do not provide information about uncertainty in a way that you can compare them, if you want, for instance, to identify which output have the greatest uncertainty. Second, through MI, the proposed method shows the impact of each input on each output. This can be further explored for determining the parameters for which it may be useful to collect more data for instance, or also as an input parameter

selection and model reduction method, for selecting the parameters that should be eliminated in case the simulation model must be simplified. Finally, for a fixed number of bins, the method also gives information about the additional benefit that the additional group of replications is providing in the reduction of the output uncertainty. This can be useful when computational power is limited and one wants to estimate the benefits/costs of running extra replications. Clearly, it is important to further investigate this to check whether there will be inflection points or not and whether a linear or quadratic function can be estimated. The proposed method is also different from Song and Nelson (2013) and Song and Nelson (2015), as their work focus on quantifying the input uncertainty when the input model are estimated from limited real-world data and how to appropriately adjust the confidence intervals due to this uncertainty.

Meanwhile, based on the results, the recommendation when using the method is: (1) to use the histogram-based method with fixed number of bins and probability density function for which a method was proposed in this work to deal with the challenges found while applying entropy measures for continuous variables. The method proposed in this work is based on data normalization, which not only eliminates some of the challenges of working with continuous variables, but also allows for easy and fair comparison among the different entropy measures. When using the probability density function to calculate the entropy measures, the suggestion is to use number of bins around 1,000 as this was the number for which the entropy measures showed the best results and that was able to fully capture the deterministic behavior. Nevertheless, based on the MI, the recommendation

would be to use a number of bins around 25 or 50. The second suggestion would be to use the theoretical normalized entropy and MI measures calculated using the histogram-based method with fixed number of bins and discrete empirical distribution. The normalization of the entropy and MI measures is recommended to eliminate the effect of the bins on the measures. These measures present the best results in capturing the deterministic behavior and for number of bins around 200 or 500, they also present entropy performance results that are as good as the ones from the probability density function. Moreover, in this case, the recommendation of the number of bins based on the entropy and MI performance would match.

As already mentioned, there are still many open questions about this topic and more experiments must be performed in different contexts in order to investigate the potential of entropy measures as a method of uncertainty quantification in simulation models. But as the results indicated, the method appears to be valuable as a method of uncertainty quantification and for identifying the inputs with the greatest impacts on the outputs. Future research in this area is invaluable and deemed necessary.

3. AN INVESTIGATION OF INFORMATION THEORY AS A METHOD FOR UNCERTAINTY QUANTIFICATION IN SIMULATION MODELS USING KERNEL METHOD, K-NEAREST NEIGHBORS, AND FUZZY-HISTOGRAM-BASED METHOD WITH STATIONARY UNIVARIATE DISTRIBUTIONS

3.1. Introduction

In 1948, Claude Shannon in his paper entitled "A Mathematical Theory of Communication" introduced the concept of entropy as a measure of information and uncertainty (Shannon, 1948). After Shannon's initial concept of entropy, many other information entropy measures have been proposed, such as Renyi's entropy, Kolmogorov-Sinai entropy, approximate entropy, sample entropy, and several others. According to Kapur (1983), these other mathematical entropies do not measure the same characteristic and their definitions have been motivated by diverse considerations. These different measures have been developed as an attempt to generalize the axioms proposed by Shannon and due to the generalization, Kapur (1983) highlighted that they may violate some of the essential properties required (or expected) from a measure of information or the underlying uncertainty.

The research studies on the area of entropy measures led to the development of the field of information theory. Since its first proposal, entropy measure has become one of the most common methods used to quantify complexity, uncertainty, and the amount of information present in real-world systems (Lacasa & Just, 2017; Xiong et al., 2017).

Despite the advantages of entropy measures for quantifying information and the relationships between variables, its practical application, especially for continuous data, is not simple due to the existing variety of entropy measures and estimators (Xiong et al., 2017). Entropy measures also require an estimate of the probability distribution of the underlying data and the method to compute the estimate without the introduction of bias in the resulting measure remains an open problem (Kinney & Atwal, 2014). Moreover, the application of entropy measures to continuous variables is not a limit of Shannon's entropy of discrete approximations and it brings some challenges and issues of interpretability as discussed earlier in section 2.4.1.

There are three main estimators discussed in the literature: the histogram-based method, the kernel-based method, and the k-nearest neighbors method (Xiong et al., 2017).

In the histogram-based method, the probability density functions (PDF) are approximated using histograms where the continuous data is divided into bins and the number of elements in each bin is counted (Dionisio et al., 2004). Selecting the bin size (or binwidth) is the main source of error. Mutual information estimates based on this binning procedure are often called naïve estimates as they may be overestimated or underestimated $I(X;Y)$ (Dionisio et al., 2004; Kinney & Atwal, 2014). The histogram-based method is computationally very efficient. In the kernel-based method, kernels are used to approximate the PDF by combining basis functions. According to Estévez et al. (2009), the quality of this approach is high, but the computational load is also significant. It is also worth pointing out that similar to the bin size selection in the histogram-based

method, in the kernel-based method the smoothing parameter is called bandwith and, consequently, the choice of the bandwidth also has an impact in the final result of the kernel-based method, as well as the kernel function chosen. Finally, in the k-nearest neighbors method entropies are estimated from KNN distances (Haeri & Ebadzadeh, 2014; Kraskov, Stögbauer, & Grassberger, 2004). Tesmer and Estévez (2004) argue that this method has an accuracy closer to the kernel-based method and it is as fast as the histogram-based method.

Haeri and Ebadzadeh (2014) proposed an adaptation to the histogram-based method, where fuzzy partitioning was used for classifying the data. According to the authors, the fuzzy partitioning approach uses a general form of fuzzy membership functions, which includes the class of crisp membership functions as a special case. The authors also argued that the method showed an average absolute error less than that of the naïve histogram-based method.

The main research question is: can entropy and mutual information measures quantify the uncertainty and, consequently, the information present in simulation models? In section 2, Shannon's entropy and mutual information measures were calculated using the histogram-based method. Their potential as measures of uncertainty quantification in simulation models were investigated through different binwidths, different normalization methods, different parameter values, and different systems. This section is also restricted to simulation models using stationary univariate distributions, because this gives a good illustrative example for which closed-form solutions are available for validation of the results.

The central contribution of this section is to extend the work of section 2, by providing an analysis of Shannon's entropy and mutual information as measures of information and uncertainty in simulation models using different estimators, namely kernel-based method, k-nearest neighbors, and fuzzy-histogram-based method. This analysis is done by: (1) continuing the discussion on the challenges of computing entropy measures for continuous variables; (2) investigating the entropy and mutual information as measures of uncertainty for different estimators and different estimators parameters (kernel function, bandwidth, fuzzy membership function, etc.); (3) investigating the measures for different normalization methods, different parameter values, and different contexts (different seeds for generating random numbers, CONWIP, and addition of travel time); (4) assessing the potential of the measures as an uncertainty quantification method in simulation model; and, (5) comparing the method when using these aforementioned different entropy estimators.

The rest of this section is organized as follows: section 3.2 provides a brief theoretical background that is important for the study being conducted. Next, section 3.3 provides an overview per entropy estimator on how entropy measures will be applied to quantify uncertainty in simulation models. Results are discussed in section 3.4, and concluding remarks and future research directions are presented in section 3.5.

3.2. Background

Shannon (1948) in his original work proposed information or uncertainty to be quantified by a new measure named entropy, as shown in Equation 48.

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \qquad \text{Equation 48}$$

Shannon's measure possesses many important properties, such as: non-negativity, it attains maximum value when the probabilities are equal, and it is consistent and additive (Kittaneh et al., 2016).

Shannon (1948) assumed, without any derivation, that the analog expression of $-\sum_{i=1}^{n} p(x_i)\, logp(x_i)$ for the continuous case, known as differential entropy, was expressed as shown in Equation 49 (Awad & Alawneh, 1987; Jaynes, 1968).

$$H'(\mathrm{X}) = -\int f(x)\, log\, f(x)\, dx \qquad \text{Equation 49}$$

This assumption without mathematical derivation resulted in the following main issues: (1) differential entropy may result in negative values, and (2) lack of invariance under linear change of variables $x \to y(x)$ (Jaynes, 1968; Kittaneh et al., 2016). According to Awad and Alawneh (1987), in Shannon's applications the issues were not significant and did not affect the final results. However, the use of $H'(X)$ leads to results that depend heavily on the choice of variables. To resolve the issues, Jaynes (1962) proposed the use of Equation 50:

$$H'(X) = -\int f(x)\, log\, \frac{f(x)}{m(x)}\, dx \qquad \text{Equation 50}$$

where $m(x)$ is a well-behaved continuous invariant measure function. In his works, Jaynes never specified an explicit form for $m(x)$ (Jaynes, 1957, 1962, 1968). Awad and Alawneh (1987) proposed the use of $m(x) = sup_{x \in R_x} f(x)$, while Kittaneh et al.

(2016) proposed the use of $m(x) = E[f(x)]$. By using $m(x) = sup_{x \in R_x} f(x)$, the differential entropy becomes positive and consistent (Kittaneh et al., 2016).

Another solution for the differential entropy was proposed by Rao et al. (2004) and it was named cumulative residual entropy (CRE). The CRE involved using the cumulative distribution of the random variable X, instead of the probability density function. However, the difficulty is that the joint CRE in this case is not defined as a natural extension of the CRE. The CRE is given by Equation 51:

$$CRE(X) = \int (1 - F(x)) log(1 - F(x))\, dx \qquad \text{Equation 51}$$

The computation of these information-theoretic measures from real-world data is challenging as it involves a two-step process, where: first, the probability mass function or the probability density function has to be estimated, and, thereafter, the entropy or mutual information can be calculated.

Several probability estimators have been proposed in the literature to address the first task and they differ in the approach used to approximate the probability function utilized in the computation of the entropy measures. The estimators are typically classified into two different groups: model-based estimators (parametric) or model-free estimators (non-parametric) (Xiong et al., 2017). Model-based estimators involve entropy measures that are calculated using functions of the parameters of the known parametric probability distribution. Model-free estimators have no fixed structure and involve entropy measures that are calculated by approximating the probability distribution directly from the data. According to Pace (1995) and Xiong et al. (2017), histogram-based method, linear

estimator, kernel estimator, and k-nearest neighbors estimator are the most common methods used for calculating entropy measures.

There are two histogram methods: the equidistant histogram method and the equiprobable histogram method. In the most common histogram method, the equidistant method, the probability estimate $\hat{f}^{hist}(x)$ is calculated by dividing the range of sampled data into n equally sized bins. The probability density function is estimated by counting the number of data points that fall into each bin and assigning to that bin a probability equal to the number of points it contains divided by the total number of data points and the binwidth, as shown in Equation 52 (Pace, 1995). To estimate the probability mass function, the number of data points that fall into each bin is counted and divided by the total number of data points. There is also the equiprobable histogram method, where rather than forming bins of equal width, one forms bins of equal mass (equal number of points in each bin) (Bonnlander & Weigend, 1994). In this case, the probability associated with each bin is given by the number of points, which is roughly equal, divided by the size of the bin.

$$\hat{f}^{hist}{}_j(x) = \frac{1}{nh}\sum_{i=1}^{n} \boldsymbol{I}\{x_i \in [t_j, t_{j+1})\} \; for \; x \in B_j, j = 1, \dots, k \qquad \text{Equation 52}$$

where h is the binwidth, $B_j = [t_j, t_{j+1})$ denotes the j^{th} bin of a total of k bins, and $\boldsymbol{I}(.)$ is the indicator function which is 1 if $x_i \in [t_j, t_{j+1})$ and 0 otherwise.

As seen from Equation 52, the histogram method is highly dependent on the choice of the binwidth and also on the choice of the origin of the bin (or the start point of the bin) (Härdle et al., 2012; Xiong et al., 2017). Another drawback of the histogram method is

that it is not a continuous function and it is not differentiable at the boundaries of the bins, which is undesirable to estimate a continuous probability density function.

The problem of choosing the start point of the bin can be alleviated by using the kernel estimate method (Härdle et al., 2012). The kernel estimate method removes the dependence on the start point of the bins by centering each of the bins at each data point rather than fixing the start point of the bins. So, one can see the procedure as counting the number of data points that fall into the interval around x. However, the issue of the binwidth has an equivalent in this method: a dependence on the choice of the bandwidth. Similar to the histogram estimate method, where the smoothing parameter was called binwidth, in the kernel estimate method the smoothing parameter is a non-negative funtion called bandwidth. Changing the bandwidth changes the shape of the kernel. A lower bandwidth means that only points close to the data are given any weight, while a larger bandwidth means that distant points also contribute to the estimate. Therefore, the resulting kernel method estimates are strongly influenced by the choice of the bandwidth: a too small bandwidth may lead to an undersmoothed kernel-density estimate (low bias, but high variance), while a too large bandwidth may lead to an oversmoothed estimate (low variance, but high bias). Hence, there is a trade-off between reducing the bias and variance of the estimator. According to Scott (2015) and Li and Racine (2007), the quality of the kernel-density estimate has been recognized to be considerably determined by the choice of the bandwidth and relatively insensitive to the choice of the kernel function. The kernel-density estimate is shown in Equation 53.

$$\hat{f}^{Kernel}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \qquad \text{Equation 53}$$

The kernel weighs information differently depending on how far they are from the value point x. $K(.)$ is the chosen kernel (weight) function and h is the bandwidth.

For the multivariate case, different approaches can be taken: (i) the use of the product kernel as in Equation 54, or (ii) pre-whitening the data by linearly transforming the data to have unit covariance matrix, estimating the density using a radial symmetric kernel, and, then, transforming it back (Scott, 2015; Silverman, 1986). According to Scott (2015), product kernels are, in practice, recommended. Silverman (1986) mentioned that by normalizing the data there is, in general, no need to consider more complicated forms of multivariate kernel density estimate.

$$\hat{f}^{Kernel}(\boldsymbol{x}) = \frac{1}{nh_1 \ldots h_d}\sum_{i=1}^{n}\left(\prod_{j=1}^{d} K\left(\frac{x - x_i}{h_j}\right)_j\right) \qquad \text{Equation 54}$$

The Epanechikov kernel function is frequently used in non-parametric estimation because it has the lowest asymptotic mean square error, but there are several other kernel functions such as Gaussian, triangular, and uniform (Scott, 2015). According to Scott (2015), well-known functions, like the Gaussian kernel, for instance, may not be the first recommendation as a kernel function due to computational overhead for computing its exponentials, its relatively inefficiency, and its infinite support.

Regarding the choice of the bandwidth, the optimal choice is the one that minimizes the MISE. This can be calculated using the Silverman's rule of thumb. For the Gaussian kernel estimator, Silverman's rule of thumb is obtained as shown in Equation 55

(Moon, Rajagopalan, & Lall, 1995; Scott, 2015). The equivalent kernel rescaling function proposed by Scott (2015) can be used to calculate the optimal bandwidth for the Epanechnikov kernel estimator. The equivalent rescaling function is shown in Equation 56.

$$h^* = \left(\frac{4}{d+2}\right)^{1/(d+4)} \sigma n^{-1/(d+4)} \qquad \text{Equation 55}$$

$$h_2{}^* = \frac{h_1{}^*}{\sigma_{K_2}} = \sqrt{5}\, h_1{}^* \approx 2.236\, h_1{}^* \qquad \text{Equation 56}$$

A drawback of the kernel-density estimate is its inability to deal with the tails of the distribution without oversmoothing the main part of the density (Silverman, 1986). This is due to the fixed smoothing parameter h that is constant and unrelated to x. When $f(x)$ is large at x more data points fall inside the interval $[x-h, x+h]$, than when $f(x)$ is small. A possible alternative for that is the KNN estimate, which automatically adapts to the amount of local information that is available (Li & Racine, 2007).

In the KNN method, the number of observations used to estimate the density is fixed by using a bandwidth that may vary with x (Li & Racine, 2007). More specifically, only the k observations nearest to x are used to estimate $f(x)$ and the greater the amount of local information, the smaller the range in which smoothing occurs. The estimate is obtained as shown in Equation 57.

$$\hat{f}^{KNN}(x) = \frac{k}{nV_k(x)} = \frac{k}{nc_d r_k^d(x)} \qquad \text{Equation 57}$$

where c_d is the volume of the unit sphere with radius $r_k^d(x)$ in d dimensions, which can be calculated using Equation 58, so that $c_1 = 2$, $c_2 = \pi$, and $c_3 = 4\pi/3$.

$$c_d = \frac{\pi^{d/2}}{\mathbb{\Gamma}\left(\frac{d+2}{2}\right)} \qquad \text{Equation 58}$$

$\mathbb{\Gamma}(.)$ is the $\mathbb{\Gamma}$ function defined by $\mathbb{\Gamma}(\propto) = \int_0^\infty t^{\propto-1}e^{-t}dt$.

Equation 57 can be rewritten as Equation 59:

$$\hat{f}^{KNN}(x) = \frac{1}{nR_k^d(x)}\sum_{i=1}^{N}\left(\frac{1}{c_d}\right)\boldsymbol{I}\left(\frac{\|x-x_i\|}{R_k^d(x)} \leq 1\right) = \frac{k}{c_d n R_k^d} \qquad \text{Equation 59}$$

where: $R_k^d(x) = R_k^d$ denotes the Euclidean distance between x and the k^{th} nearest neighbor of x among the $x_i's$. $\|.\|$ denotes the Euclidean norm, i.e., $\|x - x_i\| = \sqrt{\left(x_1 - x_{1_i}\right)^2 + \cdots + (x_d - x_{di})^2}$. Other distance metrics, such as Chebyshev could also be used. $\boldsymbol{I}(.)$ is the indicator function that ensures that only the k observations nearest to x are used to estimate $f(x)$. k/n plays a role similar to the bandwidth h for the kernel method.

A disadvantage of KNN is that it treats all the variables symmetrically, and hence, lacks the ability to smooth out irrelevant variables (Li & Racine, 2007). The method is prone to local noise and because $R_k^d(x)$ is not differentiable, the density estimate may have discontinuities. The method is also susceptible to the curse of dimensionality. One difference between the KNN and the kernel method is that the bandwidth in the first is now stochastic and, hence, the asymptotic analysis of the KNN estimator is more complex (Li & Racine, 2007).

In both kernel and KNN methods, when working with multiple variables data should be normalized before estimating the density to avoid extreme differences in the spread of the data (Li & Racine, 2007; Silverman, 1986).

According to Loquin and Strauss (2006), a modification of the traditional histogram-based method, called fuzzy-histogram method, appeared to be more robust than the histogram density estimator method. The fuzzy-histogram based method was also discussed by Arefi, Viertl, and Taheri (2012), Fajardo (2014), and Haeri and Ebadzadeh (2014). According to this method, the datapoints of X define the universe and the universe is partitioned into k fuzzy subsets A_k $(k = 1,2, \dots, p)$ based on a specific fuzzy membership function. For each subset A_k, there is a membership function $\mu_{A_k}(x)$ that defines the degree of truth of $x \in A_k$. The fuzzy-histogram density estimate is obtained as shown in Equation 60 (Loquin & Strauss, 2008).

$$\hat{f}^{fuzzy}(x) = \frac{\sum_{i=1}^{n} \mu_{A_k}(x)}{nh} \qquad \text{Equation 60}$$

According to Loquin and Strauss (2006), an interpolant, similar to the kernel method, can be used to estimate the fuzzy-histogram density as shown in Equation 61.

$$\hat{f}^{fuzzy}(x) = \frac{\sum_{i=1}^{n} \mu_{A_k} \left[\sum_{k=1}^{p} K\left(\frac{x - m_k}{h}\right)\right]}{nh} \qquad \text{Equation 61}$$

where m_k is the k^{th} node of the universe.

Similarly, for the multivariate case, Equation 62 can be used (Haeri & Ebadzadeh, 2014).

$$\hat{f}^{fuzzy}(x, y) = \frac{\sum_{i=1}^{n} \mu_{A_k \times B_k}(x, y)}{nh} \qquad \text{Equation 62}$$

where $\mu_{A_k \times B_k}(x, y)$ is the joint membership function of x and y. Based on the cardinality of fuzzy sets, we know that $\mu_{A_k \times B_k}(x, y) = \min\left(\mu_{A_k}(x), \mu_{B_k}(y)\right)$ (Cheng & Chen, 1999).

There are different choices for the membership function, such as the generalized normal function, the triangular membership function, crisp partition, and cosine partition (Haeri & Ebadzadeh, 2014). Table 29 shows the membership functions for the crisp, triangular, and cosine strong uniform fuzzy partition of the universe. For the definition of strong uniform fuzzy partition, please refer to Loquin and Strauss (2006), but the main conditions for strength and uniformity include: (i) for all $x \in \Omega = [a, b]$, the universe, $\sum_{k=1}^{p} \mu_{A_k} = 1$, and, (ii) for $k \neq p, h_k = m_{k+1} - m_k = h = constant$, so $m_k = a + (k-1)/h$.

Table 29. Membership functions of crisp, triangular, and cosine strong uniform partition of the universe.

Membership function	Crisp partition	Triangular	Cosine
$\mu_{A_1}(x)$	$\boldsymbol{I}_{\left[m_1, m_1+\frac{h}{2}\right]}(x)$	$\frac{(m_2 - x)}{h}\boldsymbol{I}_{[m_1, m_2]}(x)$	$\frac{1}{2}\left(\cos\left(\frac{\pi(x - m_1)}{h}\right) + 1\right)\boldsymbol{I}_{[m_1, m_2]}(x)$
$\mu_{A_k}(x)$	$\boldsymbol{I}_{\left[m_k-\frac{h}{2}, m_k+\frac{h}{2}\right]}(x)$	$\frac{(x - m_{k-1})}{h}\boldsymbol{I}_{[m_{k-1}, m_k]}(x) + \frac{(m_{k+1} - x)}{h}\boldsymbol{I}_{[m_k, m_{k+1}]}(x)$	$\frac{1}{2}\left(\cos\left(\frac{\pi(x - m_k)}{h}\right) + 1\right)\boldsymbol{I}_{[m_{k-1}, m_{k+1}]}(x)$
$\mu_{A_p}(x)$	$\boldsymbol{I}_{\left[m_p-\frac{h}{2}, m_p\right]}(x)$	$\frac{(x - m_{p-1})}{h}\boldsymbol{I}_{[m_{p-1}, m_p]}(x)$	$\frac{1}{2}\left(\cos\left(\frac{\pi(x - m_p)}{h}\right) + 1\right)\boldsymbol{I}_{[m_{p-1}, m_p]}(x)$
$\mu_K(x)$	$\boldsymbol{I}_{\left[-\frac{1}{2}, \frac{1}{2}\right]}(x)$	$(1 - \lvert x \rvert)\boldsymbol{I}_{[-1,1]}(x)$	$\frac{1}{2}(\cos(\pi x) + 1)\boldsymbol{I}_{[-1,1]}(x)$

Source: Loquin and Strauss (2006).

After the probability density function or the probability mass function has been estimated, the entropy and mutual information measures can be finally computed. For the calculation of the entropy measures, the evaluation of the integral in Equation 49 requires numerical integration and it is computationally inefficient, especially if $f(x)$ is a kernel density estimator (Beirlant et al., 1997; Moddemeijer, 1989). To minimize the computational burden, different estimates are available: the resubstitution estimate, the splitting data estimate, the cross-validation estimate or leave-one-out density estimate, and the sample-spacings estimates (Beirlant et al., 1997). The resubstitution estimates are shown in Equation 63 to Equation 66, where Equation 63 refers to the entropy of X, Equation 64 refers to the joint entropy of X and Y, Equation 65 refers to the conditional entropy of X given Y, and Equation 66 refers to the mutual information of X and Y (Steuer et al., 2002; Xiong et al., 2017).

$$\widehat{H}(X) = -\frac{1}{n}\sum_{i=1}^{n}\log\hat{f}(X_i) \qquad \text{Equation 63}$$

$$\widehat{H}(X,Y) = -\frac{1}{n}\sum_{i=1}^{n}\log\hat{f}(X_i,Y_i) \qquad \text{Equation 64}$$

$$\widehat{H}(X|Y) = -\frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{\hat{f}(X_i,\mathrm{Y}_i)}{\hat{f}(Y_i)}\right] \qquad \text{Equation 65}$$

$$\hat{I}(X;Y) = \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{\hat{f}(X_i,\mathrm{Y}_i)}{\hat{f}(X_i)\hat{f}(Y_i)}\right] \qquad \text{Equation 66}$$

3.3. Material and methods

In this section of the dissertation, the entropy measures will be applied to quantify the total uncertainty of the simulation outputs ($H(Y)$), the uncertainty of the simulation generated inputs ($H(X)$), and the extent of the simulation output uncertainty that can be attributed to each of the simulation generated inputs ($I(X;Y)$). Entropy measures will be calculated using the kernel method, the KNN method, and the fuzzy-histogram based method.

The experiments are run in Simio® University Enterprise Edition v 10.165 and include different system configurations by: (1) varying the traffic intensities, the number of servers, the seeds for generating random numbers, and the number of replications; (2) adding a third input (i.e., travel time); and, (3) fixing the number of customers in the system (constant work-in-progress). The model consists of a single source of arrivals, single queue, and s servers providing the same service. After being served, customers leave the system. Balking and reneging are not considered in the model. Specifically, the adequacy of the proposed method will be investigated through an $M/M/s$ and $M/G/1$ illustrative example with two simulation generated input processes and two outputs of interest, namely: $\vec{X} = [X_1 - arrival\ process, X_2 - service\ process]$ and $\vec{\hat{Y}} = [\hat{Y}_1 - average\ number\ of\ customers\ in\ the\ system, \hat{Y}_2 - average\ time\ in\ the\ system]$. The notation used in this section of the dissertation follows the $A/S/s$ Kendall's notation, where: A represents the arrival process, S the service time, and s the number of servers. M is used for memoryless distributions and G for general distributions.

3.3.1. The kernel method

For the kernel-based method, two different aspects are investigated: (1) different kernel functions, and (2) different bandwidths (different values and Silverman's rule of thumb).

Two kernel functions are investigated in this work: the Gaussian function and the Epanechnikov function. The latter is known to be the optimal kernel in the mean integrated square error sense (Silverman, 1986). The Gaussian kernel function $K(.)$ is shown in Equation 67 and the Epanechnikov kernel function is shown in Equation 68 (Li & Racine, 2007).

$$K(v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2}, -\infty \leq v \leq \infty \qquad \text{Equation 67}$$

$$K(v) = \begin{cases} \frac{3}{4\sqrt{5}}\left(1 - \frac{1}{5}v^2\right), if\ |v| < \sqrt{5} \\ 0, otherwise \end{cases} \qquad \text{Equation 68}$$

where: $v = \frac{x_i - x}{h}$.

The kernel estimator of a univariate density function can be expressed as shown in Equation 69:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K(v) \qquad \text{Equation 69}$$

where h is a smoothing parameter called bandwidth.

For the mutual information calculation, there is a need to resort to the multidimensional case. Multivariate kernel estimators can be based on the product of univariate kernel functions or on the general multivariate kernel estimator. Here, product kernel is the adopted approach, which can be estimated using Equation 54 where the same

univariate kernel is used in each dimension but with a different smoothing parameter for each dimension.

While nonparametrical kernel estimation has been recognized as fairly insensitive to the choice of kernel function, it is known that different bandwidths can lead to very different impressions of the underlying distribution (Li & Racine, 2007). Therefore, investigating the impact of different values of bandwidths on the entropy measures and on quantifying uncertainty in simulation models is important to analyze the robustness of the theory. It is known that a too small bandwidth reduces bias, but may lead to an undersmoothed kernel density estimate and increased variance. On the other hand, a too large bandwidth reduces variance, but may lead to an oversmoothed kernel density estimate and increased bias. First, different values of bandwidth ranging from values close to 0 to large values are used: 0.0001, 0.001, 0.01, 0.5, 0.2, 0.1, 1, 1.5, 5, 10, 100, 1000. Next, Silverman's rule of thumb is applied to calculate the optimal bandwidth to be used on the Gaussian kernel estimator, as shown in Equation 70 (Moon et al., 1995; Scott, 2015). According to Silverman's rule of thumb, the optimal choice of bandwidth is the one that minimizes the mean integrated square error. The equivalent kernel rescaling function proposed by Scott (2015) can be used to calculate the optimal bandwidth for the Epanechnikov kernel estimator. This rescaling function is shown in Equation 71.

$$h^* = \left(\frac{4}{d+2}\right)^{1/(d+4)} \sigma n^{-1/(d+4)} \qquad \text{Equation 70}$$

$$h_2{}^* = \frac{h_1{}^*}{\sigma_{K_2}} = \sqrt{5}\, h_1{}^* \approx 2.236\, h_1{}^* \qquad \text{Equation 71}$$

According to Silverman (1986), p. 77, as in many multivariate statistical procedures to avoid extreme differences in the spread of the data it is recommended to

pre-scale the data before attempting to estimate the kernel density. All the data is normalized prior to estimating the kernel density, as well as the KNN and the fuzzy-histogram density estimators.

3.3.2. The k-nearest neighbors method

For the k-nearest neighbors method, one aspect is investigated: (1) different values of k-nearest neighbors are used to estimate the probability density, similar to the bandwidth concept. Equation 59 is used to calculate the estimates, where k determines the number of observations that are used to estimate $f(x)$. Here, k is varied based on the number of observations available in each experiment. For the analysis of the impact of the different number of k-nearest neighbors on entropy and MI, the goal is to have different values of k that corresponds to low, medium, and high values in comparison with the amount of data available in each experiment. Table 30 shows the values of k used for this corresponding to the amount of data available. However, a different set of values of k ($k = 1,2,3,4,5,6,7,8,9$) will be used to compare the entropy with other measures of error, namely SAE, SSE, MAE, and MSE, in section 3.4.8

Table 30. Values of $\boldsymbol{k}$ used to estimate the probability density function based on the number of datapoints available.

Number of datapoints available	Different values of k investigated
10	[1, 2, 3, 4, 5, 6, 7, 8, 9]
20	[1, 3, 5, 7, 9, 11, 13, 15, 17, 19]
50	[1, 3, 5, 7, 9, 13, 19, 25, 35, 49]
100	[1, 3, 5, 7, 9, 13, 19, 25, 49, 99]
200	[1, 3, 7, 9, 19, 25, 49, 99, 150, 199]
400	[1, 3, 9, 19, 25, 49, 99, 150, 199, 399]
600	[1, 3, 9, 25, 49, 99, 150, 199, 399, 599]

800	[1, 3, 9, 25, 49, 99, 199, 399, 599, 799]
1,000	[1, 3, 9, 25, 49, 99, 199, 399, 599, 999]
1,500	[1, 3, 9, 25, 49, 99, 199, 599, 999, 1499]

3.3.3. The fuzzy-histogram based method

For the fuzzy-histogram based method, two different aspects are investigated: (1) different fuzzy membership functions, and (2) different number of fuzzy subsets (similar to the concept of bins in the histogram method).

Three membership functions are investigated in this work: the crisp function, the triangular function, and the cosine function. The assumption adopted in this section of the dissertation is of the strong uniform fuzzy partition as mentioned in section 3.2. The aforementioned functions are chosen for being commonly referred in the fuzzy literature. Regarding the number of fuzzy subsets, similar numbers adopted in the histogram-based method investigated in section 2 are being used in this section, which allows for comparison of the results. Therefore, the following different number of fuzzy partitions are investigated: $l = [2, 5, 10, 25, 50, 100, 200, 500, 1000]$. Because l is the number of partitions, there is actually a total number of $k = l + 1$ fuzzy subsets, which satisfies the condition that $k \geq 3$ (Loquin & Strauss, 2006). The nodes of each fuzzy subset are defined such that: $m_1 = \min(x_n)$, $m_p = \max(x_n)$, and $m_k = m_1 + (\mathrm{k} - 1)\mathrm{h}$, for $k \neq 1$ and $k \neq p$, where $\mathrm{h} = \frac{m_p - m_1}{p-1}$. Equation 61 is used to estimate the fuzzy histogram probability density.

3.4. Results and discussion

3.4.1. Challenges encountered while applying entropy measures for continuous variables and method proposed to overcome the issues

In section 2.4.1, the challenges encountered while applying entropy measures for continuous variables were discussed. These challenges go beyond requiring an estimate of the probability distribution of the underlying data.

The first solution proposed in the literature to tackle these challenges is to approximate the differential entropy by calculating the discretized entropy and adjusting it with a correction dependent on the binwidth. This solution eliminates the issue of entropy being infinity, but does not eliminate the issue of the entropy being a negative value.

In section 2, the entropy was calculated based on the histogram-method and a data normalization method, that guaranteed the data was always between 0 and 1, was proposed to handle the challenges of calculating entropy for continuous variables. The procedure resulted in non-negative entropy values and also solved some issues of interpretability in the continuous case. However, the method could only be implemented when using fixed number of bins. When formulas that calculate optimum number of bins or entropy estimators that do not use number of bins are used, the method cannot be applied because the binwidth is not a function of the data range.

Another approach commonly cited in the literature was proposed by Jaynes (1962). In Jaynes' approach, differential entropy is calculated as $H(X) = -\int p(x) \log\left[\frac{p(x)}{m(x)}\right] dx$ and, hence, it is important to define the function $m(x)$. Some authors suggest to use $m(x) = sup\left[\hat{f}(x)\right]$ or $m(x) = E\left[\hat{f}(x)\right]$ (Awad & Alawneh, 1987; Kittaneh et al., 2016).

Clearly, $m(x) = E\left[\hat{f}(x)\right]$ is not a good choice as $\hat{f}(x)$ may be greater than $E\left[\hat{f}(x)\right]$. In this case, $\frac{\hat{f}(x)}{m(x)} > 1 \Rightarrow log_2 \frac{\hat{f}(x)}{m(x)} > 0$ and, $H(X)$ may be negative. Here, two other functions are proposed: $m(x) = \hat{f}(x)^2\left(\hat{f}(x) - 1\right)^{-1}$ and $m(x) = \hat{f}(x)\left(1 + \hat{f}(x)\right)$. Table 31 shows a discussion of the challenges and issues encountered when using each of these functions and Table 32 summarizes these issues per proposed $m(x)$.

Table 31. Issues and challenges encountered when using the different proposed $\boldsymbol{m(x)}$.

$\boldsymbol{m(x)}$	**Challenge or issue**
$m(x) = sup[f(x)]$	- When events are all equiprobable (or all probability density function values are equal) the differential entropy is 0, but the discrete entropy is maximum. - When there is a mix of events that are equiprobable and events that will surely not occur (the probability density function values are equal or 0, respectively) the differential entropy is 0, but the discrete entropy is maximum. - An event will surely occur (certainty): there is an open question regarding how to represent certainty in the nonparametric continuous case. If $\hat{f}_i(x) = \sup \hat{f}(x)$ is considered the certainty, the differential entropy is 0, same as the discrete entropy. However, one could interpret that certainty does not exist in the continuous case as $P(X = x) = 0$.
$m(x) = \hat{f}(x)^2\left(\hat{f}(x) - 1\right)^{-1}$	- When probability density function is between 0 and 1, the proposed solution will be negative and, hence, the logarithm does not exist. - An event will surely occur: similar to above, it is an open question. If one considers that certainty does not exist in the continuous, an issue does not exist. On the other hand, if certainty exists, what would be considered the certainty value for which the entropy should be equal to 0? $\hat{f}_i(x) = \sup \hat{f}(x)$?
$m(x) = \hat{f}(x)\left(1 + \hat{f}(x)\right)$	- When probability density function is 0, the proposed solution is 1, which could be an issue for the scenarios

	with events that will surely not occur. However, $\log_2(1) = 0$ and, hence, no issue exists. - An event will surely occur: similar to above, it is an open question. If one considers that certainty does not exist in the continuous, an issue does not exist. On the other hand, if certainty exists, what would be considered the certainty value for which the entropy should be equal to 0? $\hat{f}_i(x) = \sup \hat{f}(x)$?

Table 32. Summary of the challenges encountered per proposed $\boldsymbol{m(x)}$.

Challenge or issue / $\boldsymbol{m(x)} =$	$\boldsymbol{sup[f(x)]}$	$\boldsymbol{\hat{f}(x)^2\left(\hat{f}(x)-1\right)^{-1}}$	$\boldsymbol{\hat{f}(x)\left(1+\hat{f}(x)\right)}$
Events equiprobable	√	X	X
Events equiprobable + events will surely not occur	√	X	X
All events will surely not occur	X	X	X
Event will surely occur (certainty)	X	*	*
Probability density function between 0 and 1	X	√	X

Legend:
√ challenge is encountered
X challenge is not encountered
* challenge may be encountered

Based on the challenges discussed on Table 31 and Table 32, one can see that $m(x) = \hat{f}(x)\left(1 + \hat{f}(x)\right)$ is the function that shows most potential to be used in Jaynes' method and, therefore, it will be the function used here. The entropy measures will be calculated using the approach proposed by Jaynes (1962) in order to allow for the calculation of the measures regardless of the choice of the bin and the choice of the density estimation method.

3.4.2. Challenge encountered when using the kernel method

Another challenge was encountered when using the kernel method in the multivariate case, that is, more specifically when calculating the MI information measures. From the experiments and based on the product kernel multivariate formula given by Equation 54, we know that $K\left(\frac{x-x_i}{h_j}\right)_j \geq 0 \Rightarrow \prod_{j=1}^{d} K\left(\frac{x-x_i}{h_j}\right)_j \geq 0$ $\Rightarrow \sum_{i=1}^{n}\left(\prod_{j=1}^{d} K\left(\frac{x-x_i}{h_j}\right)_j\right) \geq 0$. When $h_1, \ldots, h_d$ are very small, that is $h_{\mathrm{j}} < 0$, the product $h_1 \ldots h_d$ becomes very small. In this case, $\frac{1}{nh_1 \ldots h_d}$ becomes very large and $\hat{f}^{Kernel}(\boldsymbol{x})$ also becomes very large. Given the normalization $m(x)$ adopted in this work, the resulting joint $\hat{f}^{Kernel}(\boldsymbol{x})$ after normalization is very small.

MI can be approximated as given by Equation 66. Because the kernel density estimates $\hat{f}(X_i)$ and $\hat{f}(Y_i)$ were calculated using a single bandwidth, there was not an interaction among bandwidths. In this case, $\hat{f}(X_i)\hat{f}(Y_i) > \hat{f}(X_i, \mathrm{Y}_i) \Rightarrow \frac{\hat{f}(X_i,\mathrm{Y}_i)}{\hat{f}(X_i)\hat{f}(Y_i)} < 1$ $\Rightarrow log\left[\frac{\hat{f}(X_i,\mathrm{Y}_i)}{\hat{f}(X_i)\hat{f}(Y_i)}\right] < 0 \Rightarrow \frac{1}{n}\sum_{i=1}^{n} \log\left[\frac{\hat{f}(X_i,\mathrm{Y}_i)}{\hat{f}(X_i)\hat{f}(Y_i)}\right] < 0$, which is theoretically incorrect and indicates an issue.

According to Silverman (1986), in certain circumstances it may be more appropriate to use a single smoothing parameter h, a vector of smoothing parameters, or even a matrix of shrinking coefficients. The author reinforces that a matrix of smoothing coefficients may be more adequate when the spread in the data points is much greater in one of the coordinate directions than the others, while a single smoothing parameter is appropriate when each data point is scaled equally. The author also restates his recommendation of pre-scaling the data to avoid extreme differences in the spread of the

data in the various coordinate directions. Although Silverman (1986) mentioned about a single smoothing parameter, he still considered raising it to the power of the dimension of the data.

All the data used in this dissertation is being pre-scaled to avoid differences in the spread of the data. Due to the issue encountered while calculating the MI using such data, a solution proposed here is to calculate the product kernel using Equation 72. This guarantees that the interaction among the smoothing parameters is not causing issues in the multivariate case and, consequently, in the MI calculations.

$$\hat{f}^{Kernel}(\boldsymbol{x}) = \frac{1}{n\sqrt[d]{h_1 \dots h_d}} \sum_{i=1}^{n} \left(\prod_{j=1}^{d} K\left(\frac{x_i - x_{ij}}{h_j} \right) \right) \qquad \text{Equation 72}$$

3.4.3. The impact of different kernel functions and different bandwidth values on entropy and mutual information measures calculated using kernel method

The discussion in this and the following sections are based on the results of the experiments, which were detailed in section 3.3 and listed in Table 66 of the Appendix. For simplification, the experiments are referred by their numbers. In summary, the experiments consist of $M/M/s$, $M/G/1$, and $G/G/s$ queue models with different traffic intensities and number of replications and where different seeds were used to generate the random numbers. An additional third input (i.e., travel time) and CONWIP configurations were also investigated. Based on the results of the experiments, entropy and MI measures were calculated in three different ways: (1) using the kernel method; (2) using the k-nearest neighbors method; and, (3) using the fuzzy histogram-based method.

For the kernel method, two different aspects are of interest: (1) the impact of different kernel functions, and (2) the impact of different bandwidths on the entropy and MI measures.

The analysis of Figure 30 shows that, different from the histogram-based method with fixed number of bins and probability density function, the entropy and MI measures tend to increase with the decrease in the bandwidth for the same number of replications regardless of the kernel function being used. This behavior is similar to the observed results of the histogram-based method with fixed number of bins and discrete empirical distribution before normalization, but opposite to the one observed when histogram-based method with fixed number bins and probability density function method used to calculate the measures. If the bandwidth is interpreted as $\frac{1}{number\ of\ bins}$ as it is the binwidth in the histogram-based method, this could indicate the need for normalization of the entropy and MI measures to eliminate the impact of the bins on the measures as suggested in the literature for the histogram-based method with discrete variables. By looking at the histogram-based method probability density estimator and the kernel method probability density estimator, one could expect the binwidth (or bandwidth) to have the same impact on the entropy and MI measures. That is, according to the histogram-based method the probability density is estimated by $\hat{f}^{hist}{}_j(x) = \frac{1}{nh}\sum_{i=1}^{n} \mathbf{I}\{x_i \in [t_j, t_{j+1})\}\ for\ x_j \in B_j, j = 1, \dots, k$, and according to the kernel method the probability density is estimated by $\hat{f}^{Kernel}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right)$. Therefore, one can see that the binwidth (or bandwidth) h is being considered as a similar divisor in both estimators, although h also plays a role

within the kernel function in the kernel estimator. However, the first difference between these two estimators is that $\sum_{i=1}^{n} \mathbf{I}\{x_i \in [t_j, t_{j+1})\} \in \mathbb{Z}$ and $\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \in \mathbb{R}[0,1]$. This difference is not enough to explain why the results of the two estimators are different in terms of changes in the binwidth. The approach proposed in this section of the dissertation to calculate entropy and similarly to calculate MI is based on the work of Jaynes' and also considers the entropy and MI approximation, which leads to: $\widehat{H}(X) = -\frac{1}{N}\sum_{j=1}^{N} \log \hat{f}(X_j) = -\frac{1}{N}\sum_{j=1}^{N} \log\left[\frac{\hat{f}(x)}{m(x)}\right]$, where $m(x) = \hat{f}(x)(1 + \hat{f}(x))$. When histogram with fixed number of bins is the method being used N is equal to the number of bins, which, in turn, determines the binwidth. When the kernel method is being considered, $N = n$ which is the number of datapoints. Therefore, one can see that although the binwidth may affect the histogram and kernel probability density estimators in the same way, it does not affect the entropy estimation likewise because with the increase in the number of bins (or decrease in the binwidth), the divisor of the entropy approximation in the histogram-based method is increasing and, hence, the entropy measure is decreasing, which is the opposite of what is observed in the kernel method.

Based on the experiments performed and the range of bandwidths experimented, for low values of bandwidth (bandwidth less than or equal to 0.01 for the experiments performed in this study) the entropy and MI measures decrease with the increase of the number of replications. For mid-range values of bandwidth (bandwidth between 0.1 and 0.5) the entropy and MI measures increase with the increase of the number of replications. Finally, for large values of bandwidth (bandwidth greater than or equal to 1), the entropy

and MI measures are approximately constant with the increase of the number of replications. This may reflect the impact of the bandwidth in undersmoothing or oversmoothing the density estimation. With larger bandwidths, there is an increased risk of oversmoothing, which is potentially being captured by the entropy and MI measures as the measures become approximately equal and close to 0 regardless of the number of replications used. This can be explained by how the probability density is estimated using the kernel method. Using the kernel method, the probability density is estimated using $\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K(v)$. For the experiments considered in this section of the dissertation, n varied from 10 to 1500. More specifically, $n = 10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500$. Based on the approach proposed in this dissertation, we have: $\widehat{H}(X) = -\frac{1}{n}\sum_{j=1}^{n} \log\left[\frac{\hat{f}(x)}{\hat{f}(x)\left(1+\hat{f}(x)\right)}\right] = -\frac{1}{n}\sum_{j=1}^{n} \log\left[\frac{1}{\left(1+\hat{f}(x)\right)}\right] = -\frac{1}{n}\sum_{j=1}^{n} \log\left[\frac{1}{\left(1+\frac{1}{nh}\sum_{i=1}^{n} K(v)\right)}\right]$. When h is very small ($h = 0.0001, 0.001, 0.01$), for the experiments under investigation $\frac{1}{nh}\sum_{i=1}^{n} K(v)$ will be large, $\frac{1}{\left(1+\frac{1}{nh}\sum_{i=1}^{n} K(v)\right)}$ will be a small decimal number and, hence, $\log\left[\frac{1}{\left(1+\frac{1}{nh}\sum_{i=1}^{n} K(v)\right)}\right]$ will be a greater negative number. However, for larger number of replications, $\frac{1}{nh}\sum_{i=1}^{n} K(v)$ will be smaller, because nh is larger and, hence, $\frac{1}{\left(1+\frac{1}{nh}\sum_{i=1}^{n} K(v)\right)}$ will be a larger decimal number and $\log\left[\frac{1}{\left(1+\frac{1}{nh}\sum_{i=1}^{n} K(v)\right)}\right]$ a smaller negative number. Consequently, for small number of replications, the entropy will be larger than for large number of replications. It is worth noting that another

important explanation and condition for this to happen is the fact that with low values of bandwidth, only the own datapoint being evaluated (or datapoints very close to it in value) end up being inside the band of the kernel function and, consequently, adding any information to the calculation. Therefore, only one or few data points are adding the same (or close) amount of information in every calculation. Because this information is being averaged by n, which depends on the number of replications, when n is larger the entropy will be smaller. On the other hand, for mid-range values of bandwidth, in general, all the datapoints will be adding information. Because of that, $\sum_{i=1}^{n} K(v)$ will be larger for larger number of replications than for smaller number of replications and this will tend to compensate the fact that nh also increases with the increase in the number of replications. This way, for mid-range values of bandwidth, $\frac{1}{nh}\sum_{i=1}^{n} K(v)$ will tend to be larger for larger number of replications and, hence, $\log\left[\frac{1}{\left(1+\frac{1}{nh}\sum_{i=1}^{n} K(v)\right)}\right]$ will be a greater negative number and the entropy will be larger for larger number of replications than for smaller number of replications. It is important noting that this is considering the data of the experiments performed here and also the fact that all the data was normalized between 0 and 1. Finally, the other extreme is to consider large values of bandwidth (e.g., $h = 10, 100, 1000$). With large values of bandwidth two different things occur: (i) first, nh becomes so large that $\frac{1}{nh}\sum_{i=1}^{n} K(v)$ will tend to go to 0 and, hence, entropy will tend to go to 0 too, (ii) second, before nh becomes so large, one can already see that when h is equal to 1 or 1.5, the entropy is already constant over the number of replications. The reason is that when h is large all datapoints are inside the band of the kernel function and, consequently,

contributing with information. However, the information each datapoint contributes is exactly the same. Because the information is the same, regardless of the number of replications, the final average information will be the same. In other words, with large values of bandwidth, $K(v)$ tend to be equal among the different i. Therefore, $\frac{1}{nh}\sum_{i=1}^{n} K(v)$ will be equal (or approximately equal) regardless of the number of replications and, hence, the entropy will be constant over the number of replications.

Figure 31 shows the same results as Figure 30 but only for bandwidths greater than or equal to 0.1. In Figure 31 it is possible to better perceive the observations just described. Another important observation is that the amplitude of the change is higher for lower values of the bandwidth. When the bandwidth increases the amplitude of the change becomes smaller until there is almost no change as already mentioned. Similar observations can be made whether the normal or the Epanechnikov kernel functions were used, which goes in direction with the literature that recognizes the kernel estimation is fairly insensitive to the choice of the kernel function but not to different bandwidths.

When Silverman's rule of thumb was used to calculate the bandwidth, the bandwidth is dependent on the data and, hence, varied based on that. In this case, plotting the entropy and MI measures per Silverman's bandwidth would not be very useful. Therefore, the average entropy and MI measures over all the experiments was plotted, instead per Silverman's bandwidth, as shown in Figure 32. From Figure 32 it is not possible to observe how the entropy and MI measures behave for the different Silverman's bandwidths. However, it is still possible to observe that when using Silverman's rule of thumb, the entropy and MI measures still tend to increase with the increase in the number

of replications, as they did when calculated using the histogram-based method or when using the kernel method with mid-range values of bandwidth.

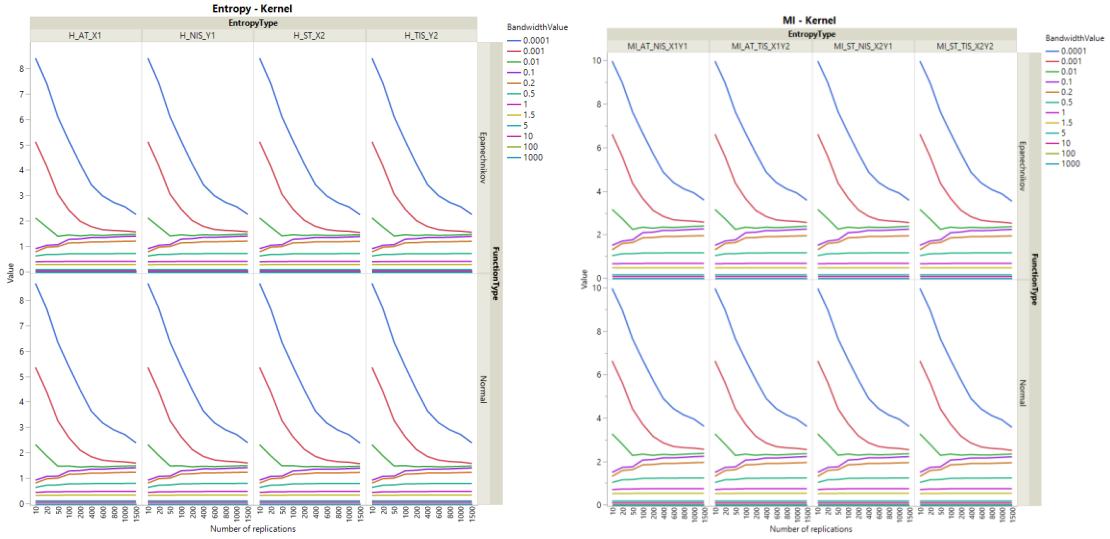

Figure 30. Average of entropy and MI measures per different bandwidths using kernel method with different kernel functions (experiments #1 to #350).

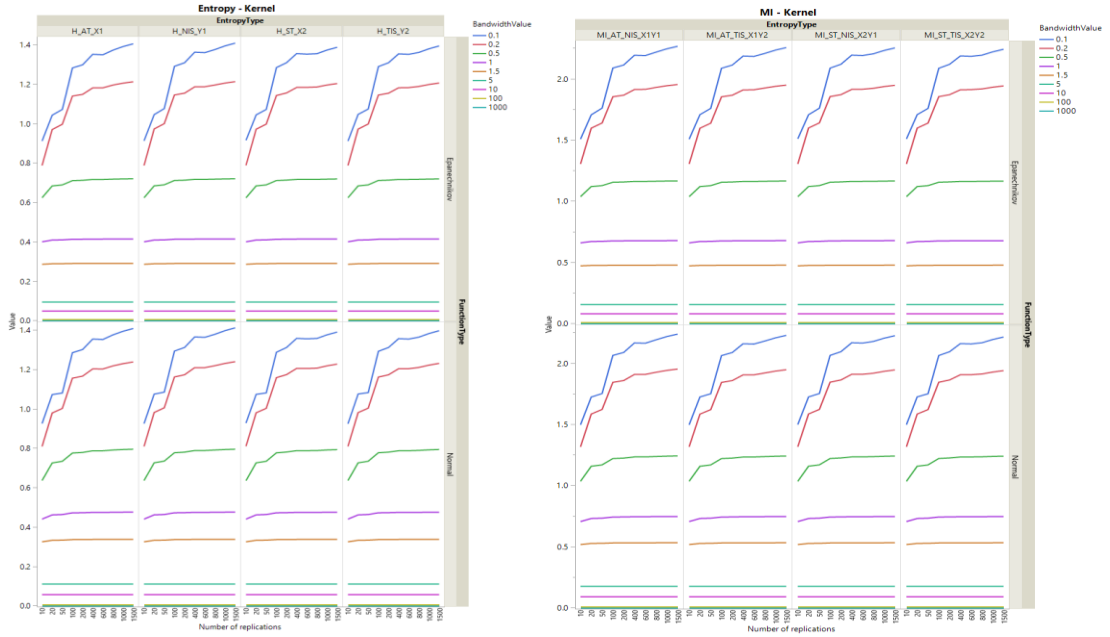

Figure 31. Average of entropy and MI measures per different bandwidths (bandwidth greater than or equal to 0.1) using kernel method with different kernel functions (experiments #1 to #350).

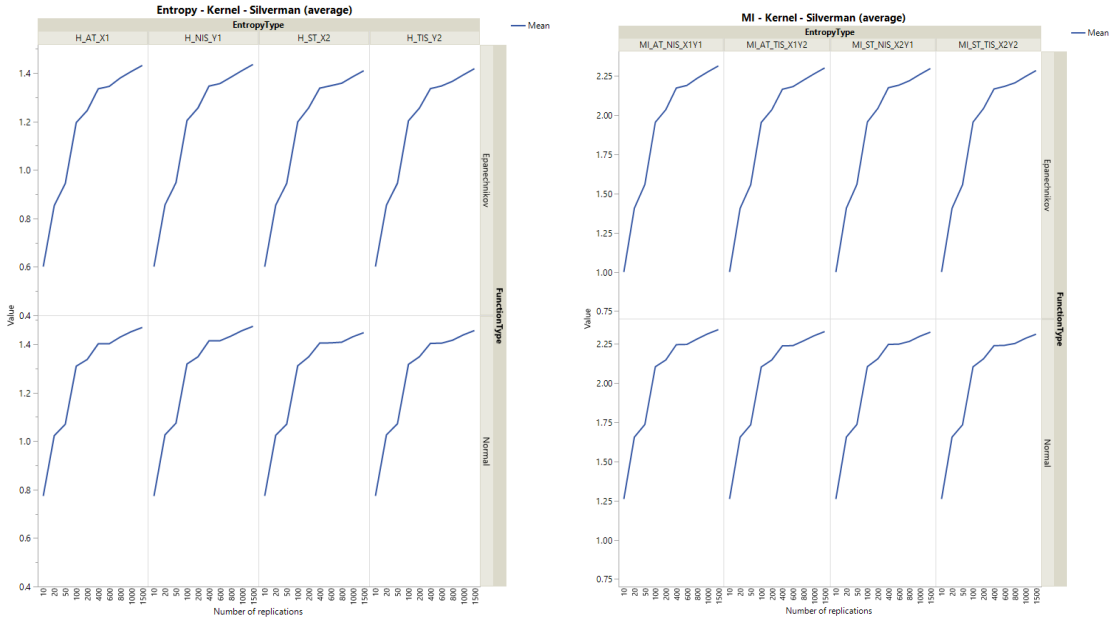

Figure 32. Average of entropy and MI measures using kernel method and Silverman's rule of thumb with different kernel functions (experiments #1 to #350).

3.4.4. The impact of different number of k-nearest neighbors (k) on entropy and mutual information measures calculated using KNN method

As mentioned by Gutierrez-Osuna (2020), the general expression for non-parametric density estimation is given by $p(x) \cong {}^{k}/_{nV}$, where k is the number of examples inside the volume V, where V is the volume surrounding the data vector x, and n is the total number of examples. V is simply the binwidth or bandwidth h. In the KNN method, k is fixed and V is determined from the data. So, for the KNN method, there is one aspect of

interest: (1) the impact of different values of k-nearest neighbors on the entropy and MI measures.

The analysis of Figure 33 shows that the entropy and MI measures tend to decrease with the increase in the number of k-nearest neighbors used to calculate the measures. By increasing the number of k-nearest neighbors used to calculate the measures, one is using more data in the volume surrounding the data point to estimate the probability density, which is translated in the method by using a larger ordered value to estimate the probability density. In the KNN method, the probability density is estimated by $\hat{f}^{KNN}(x) = \frac{k}{nV_k(x)}$. Because k is larger, $\hat{f}(x)$ is larger, $\frac{1}{\left(1+\hat{f}(x)\right)}$ is smaller, and entropy will be larger. Therefore, the fact that the entropy and MI measures tend to decrease with the increase in the number of k agrees with the fact that the entropy and MI measures tend to decrease with the decrease in the binwidth (or increase in the number of bins) when using the histogram-based method with fixed number of bins and probability density function or the histogram-based method with fixed number of bins and discrete empirical distribution after normalization, although for different reasons.

Based on the experiments performed and the different values of k experimented, the entropy and MI measures increase with the increase of the number of replications for the same value of k. The exception occurs for very low values of k (when k is less than or equal to 3, as can be seen in Figure 33 and Figure 34). The increase of the entropy and MI measures with the increase in the number of the replications was also observed in the histogram-based method with probability density function and for most of the cases when using the

histogram-based method with discrete empirical distribution regardless or not normalization was applied.

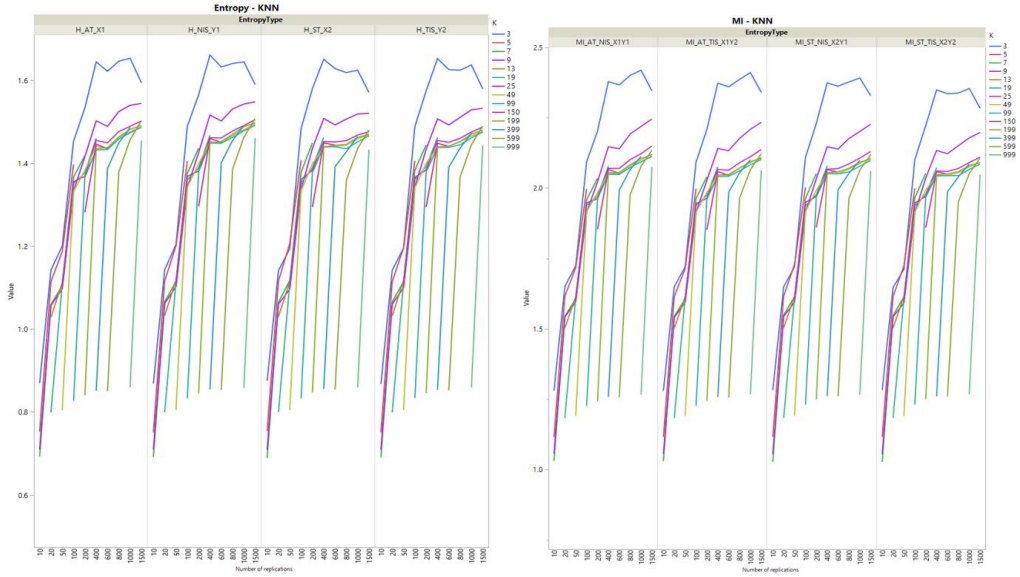

Figure 33. Average of entropy and MI measures per different values of k-nearest neighbors using KNN method (experiments #1 to #350).

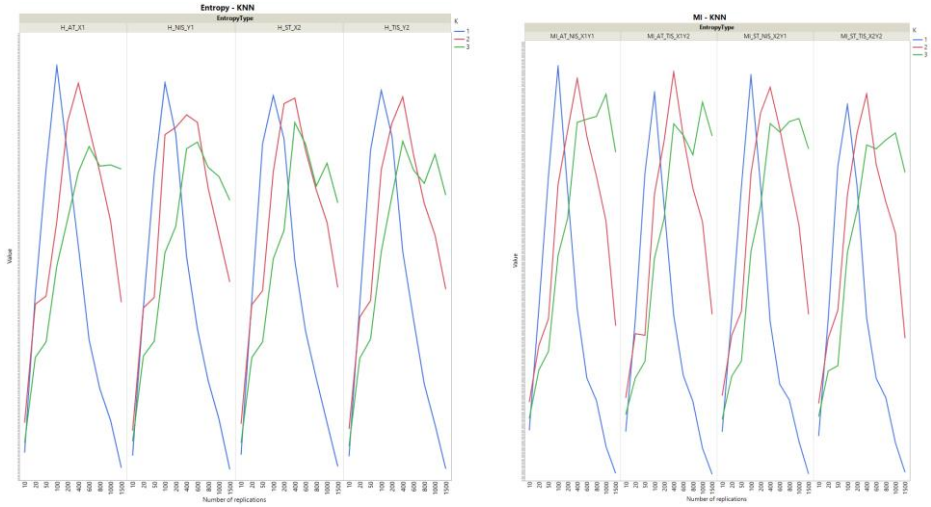

Figure 34. Average of entropy and MI measures per low values of k-nearest neighbors ($\boldsymbol{k} = \mathbf{1}, \mathbf{2}$, and $\mathbf{3}$) using KNN method (experiments #1 to #350).

3.4.5. The impact of different fuzzy membership functions and different number of fuzzy subsets on entropy and mutual information measures calculated using fuzzy histogram method

For the fuzzy-histogram based method, there are two different aspects of interest: (1) the impact of different fuzzy membership functions, and (2) the impact of different number of fuzzy subsets on the entropy and MI measures.

Figure 35 shows that, similar to the histogram-based method with fixed number of bins and probability density function, the entropy and MI measures tend to decrease with the increase in the number of fuzzy subsets for the same number of replications regardless of the fuzzy membership function. This occurs especially for the entropy measures rather than the MI measures and when larger number of fuzzy subsets are considered. The fact that the entropy and MI measures tend to decrease with the increase in the number of fuzzy subsets is similar to the results from the histogram-based method with fixed number of bins and probability density function and also the histogram-based method with fixed number of bins and discrete empirical distribution after normalization. The number of fuzzy subsets can be seen as the number of bins in the histogram-based method and this is why the behavior is similar to the one observed in the histogram-based method, because in the entropy approximation the number of fuzzy subsets is used as divisor to approximate the entropy as it is the number of bins in the histogram-based method.

Based on the experiments performed and the different number of fuzzy subsets considered, one can see that the entropy and MI measures tend to increase with the

increase in the number of replications for number of fuzzy subsets greater than or equal to 25. This behavior was also observed from other methods such as the KNN and the histogram-based method with probability density function.

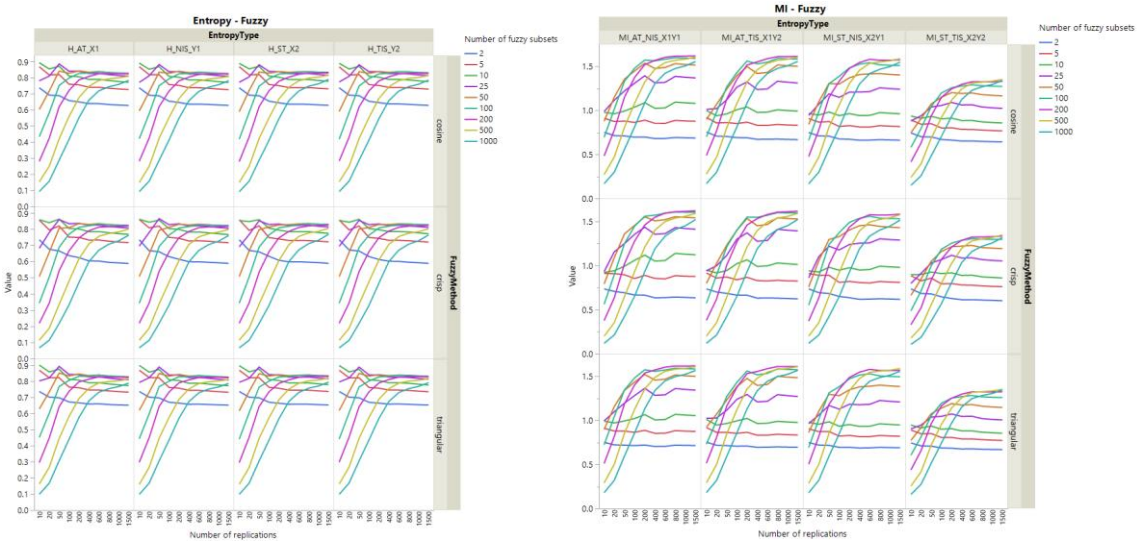

Figure 35. Average of entropy and MI measures per different values of fuzzy subsets and fuzzy membership functions using fuzzy-histogram based method (experiments #1 to #350).

3.4.6. The impact of different traffic intensities, different seeds, different parameter values, and different systems on entropy and mutual information measures based on the method used

As it was done when entropy and MI measures were calculated using the histogram-based method, the appropriateness of the measures as an uncertainty quantification method in simulation models will be investigated by discussing the impacts

of different traffic intensities, different seeds, different parameter values, and different systems on the measures.

In the queue example used in this study, it is known that the uncertainty of the input X_1 must be equal among the different traffic intensities because the same input model and fixed seed were used in the simulation model. For all the methods used (kernel method with different values of bandwidth or with Silverman bandwidth, KNN method, or fuzzy-histogram based method with different membership functions), the entropy of X_1 was equal among the different traffic intensities which indicates that the entropy measure is possibly accurately measuring the information or uncertainty of X_1. This result can be seen in Figure 36 to Figure 41.

Although a fixed seed was also used for the input X_2, it is not correct to expect that X_2 should also have equal entropy among different traffic intensities because changes in traffic intensities were modeled by changing the capacity of the only existing server, instead of creating or eliminating servers. Therefore, even though the seed of the service time input, X_2, is fixed, the generated input X_2 are different and, thus, its entropy should not remain the same among the different traffic intensities. Reviewing Figure 36 to Figure 41, one can observe that the entropy of X_2 was also able to capture some differences among the different traffic intensities for all the methods considered.

Finally, one would also expect the reduction in uncertainty in the output provided by the input to be different in a high traffic intensity than in a low traffic intensity system. That is, the MI measures should be able to capture differences among the different traffic

intensities as it did in Figure 36 to Figure 41, which indicates that the MI measure is correctly measuring the information or uncertainty of the simulation generated inputs.

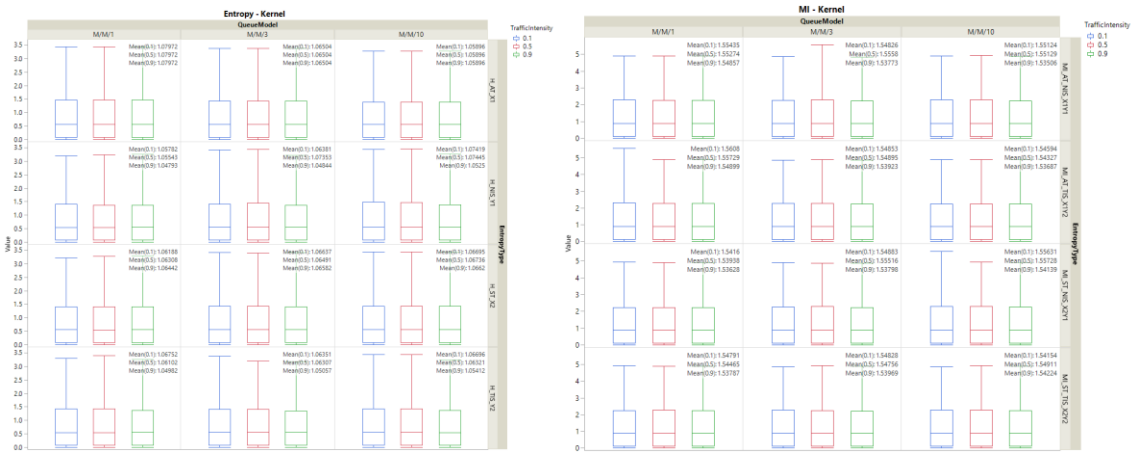

Figure 36. Entropy and MI measures per queue model per traffic-intensity using kernel method with different values of bandwidth (experiments #1 to #90).

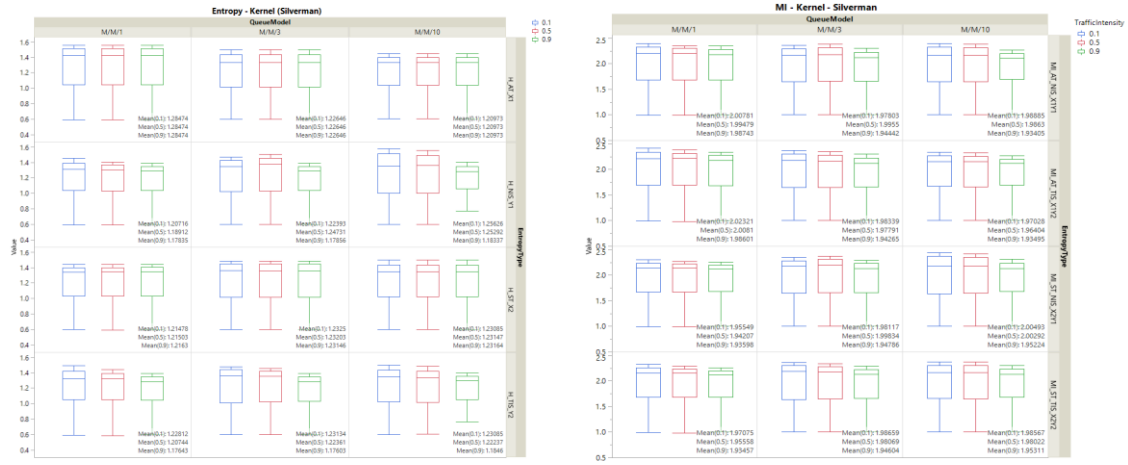

Figure 37. Entropy and MI measures per queue model per traffic-intensity using kernel method with Silverman bandwidth (experiments #1 to #90).

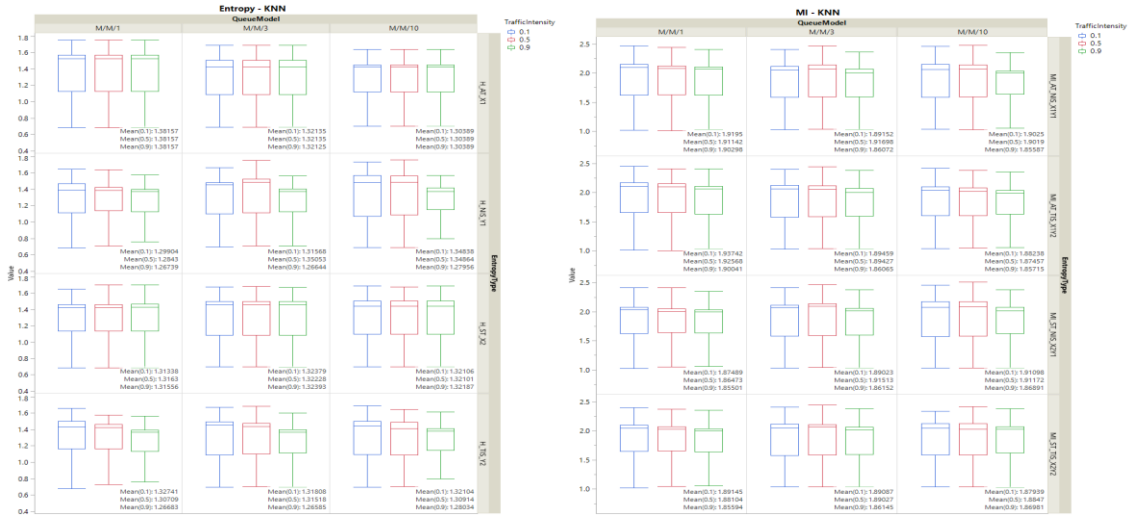

Figure 38. Entropy and MI measures per queue model per traffic-intensity using KNN method with different values of k-nearest neighbors (experiments #1 to #90).

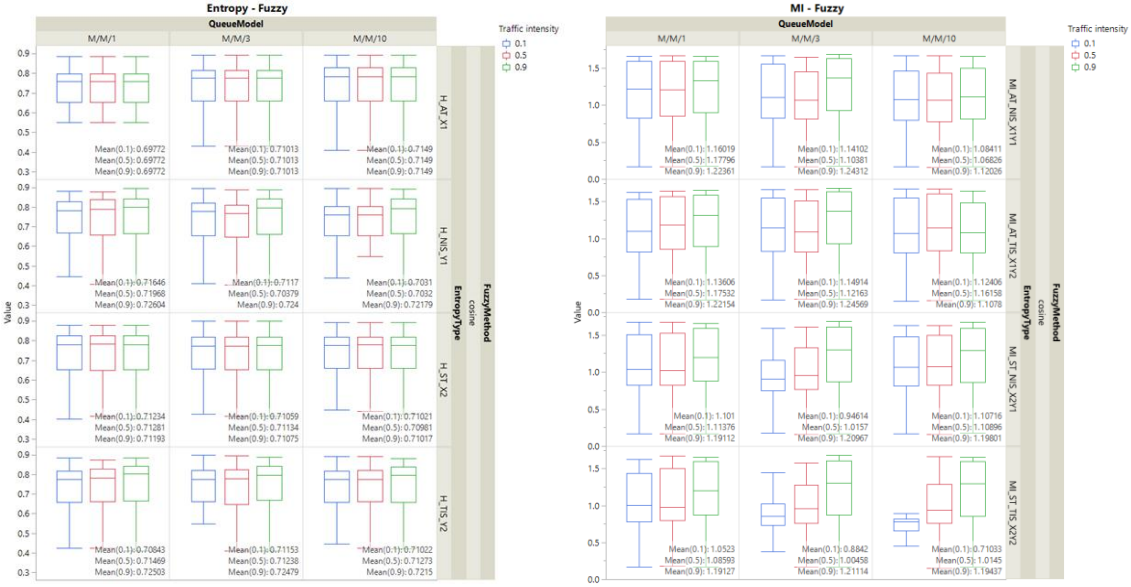

Figure 39. Entropy and MI measures per queue model per traffic-intensity using fuzzy-histogram based method with different values of fuzzy subsets and cosine fuzzy membership function (experiments #1 to #90).

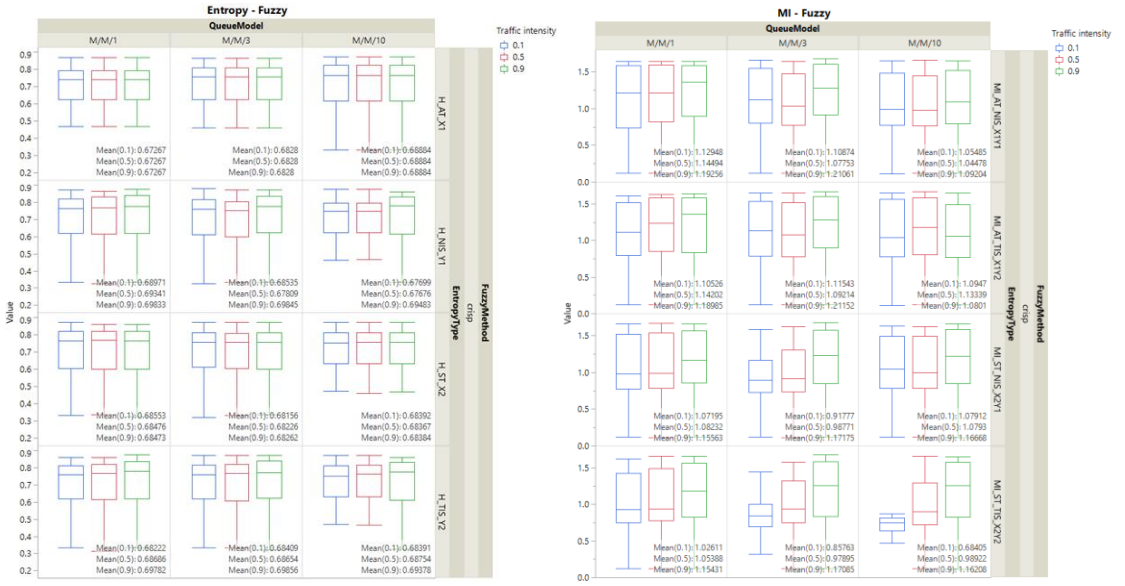

Figure 40. Entropy and MI measures per queue model per traffic-intensity using fuzzy-histogram based method with different values of fuzzy subsets and crisp fuzzy membership function (experiments #1 to #90).

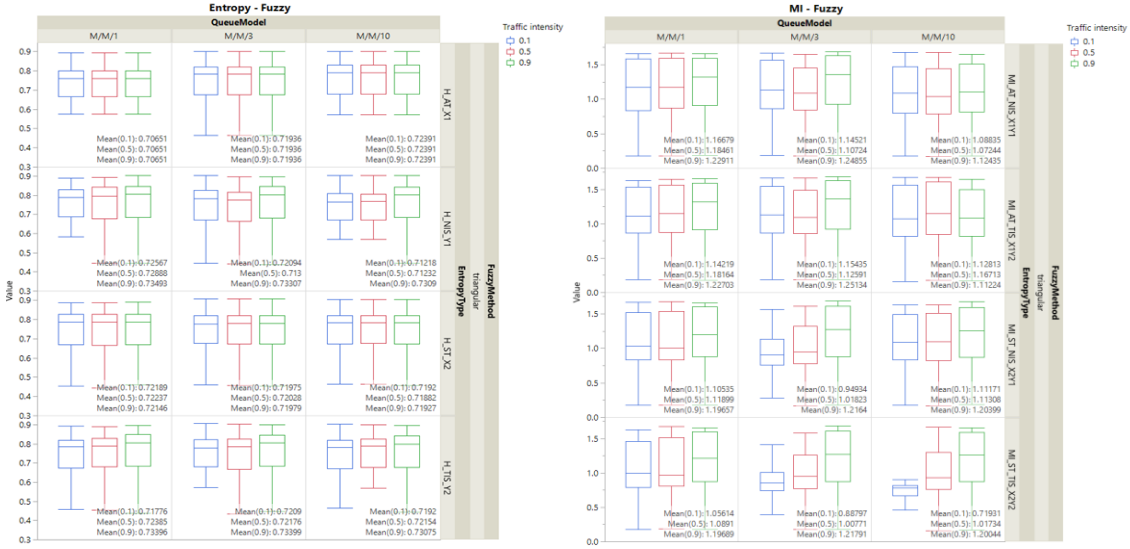

Figure 41. Entropy and MI measures per queue model per traffic-intensity using fuzzy-histogram based method with different values of fuzzy subsets and triangular fuzzy membership function (experiments #1 to #90).

Another aspect investigated was how using different seeds would impact the entropy and MI as measures of uncertainty quantification. As shown in Table 66, experiments #351 to #440 correspond to the initial experiments numbered #1 to #90, but with a different seed (here, named "seed 2", the specific seed number used is shown in Table 66). That is, the same experiment configuration was kept: same interarrival time, same service time, same number of servers, but different seeds were used for the parameters. Similarly, experiments #471 to #560 correspond to the initial experiments numbered #1 to #90 but with another seed (here, named "seed 3").

As shown in Figure 42 to Figure 48, the entropy measures and mutual information vary based on the seed used. For different seeds, there are different uncertainties. Therefore, one should expect different entropy values for the inputs, different entropy values for the outputs, and hence different MI values as the results in Figure 42 to Figure 48 show. From Figure 42 to Figure 48 one can also see that regardless of the seed and the method used to calculate the entropy measure, the entropy of X_1 is equal among the different traffic intensities as expected.

Another interesting observation that can be made from Figure 43 to Figure 48 and that is also similar from the results observed with the histogram-based method with fixed number of bins and probability density function is that when using "seed 3" the entropy of $\hat{Y}_1$ and $\hat{Y}_2$ present the greatest variability among the experiments and, a few times, even the highest values. This can be used as an indication that the seeds used as "seed 3" are likely not a good combination to be used in the experiments as they are resulting in output uncertainties that are somehow more different than the ones observed with the other seeds.

This is possibly due to some interaction between the seeds used for the simulation generated inputs X_1 and X_2 and it shows that entropy can be used to identify the adequacy of seeds in simulation studies. The same observation cannot be made from Figure 42 because the graph shows the average of the results for all the values of bandwidth experimented. As previously discussed, for larger values of bandwidth, the entropy measures tend to go to 0 and, hence, the observation being discussed here cannot be made. However, when only lower values of bandwidth are considered (e.g., bandwidth equal to 0.0001, 0.001, and 0.01), as shown in Figure 43, the same observation regarding the impact of "seed 3" on the entropy of $\hat{Y}_1$ and $\hat{Y}_2$ can be made for the kernel method with different values of bandwidth.

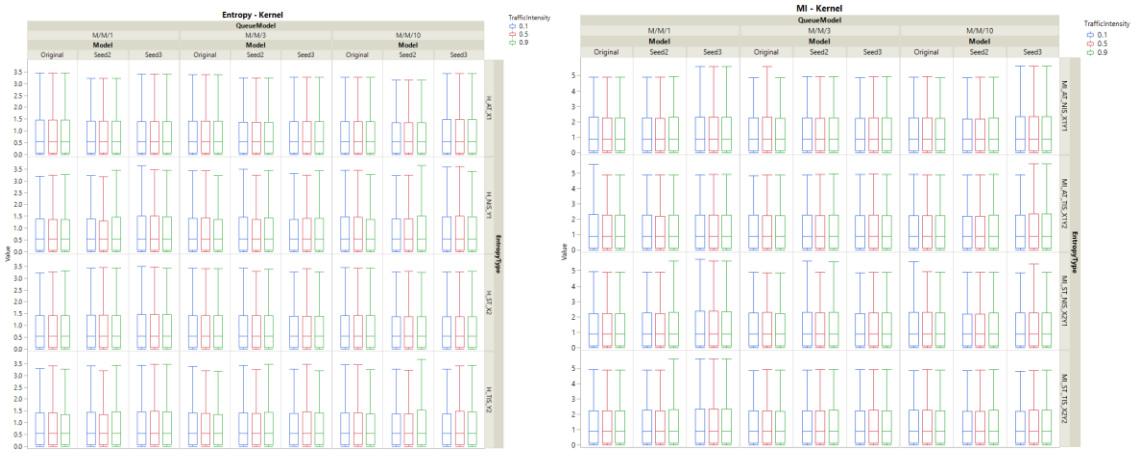

Figure 42. Entropy and MI measures per queue model per traffic-intensity per seed using kernel method with different values of bandwidth.

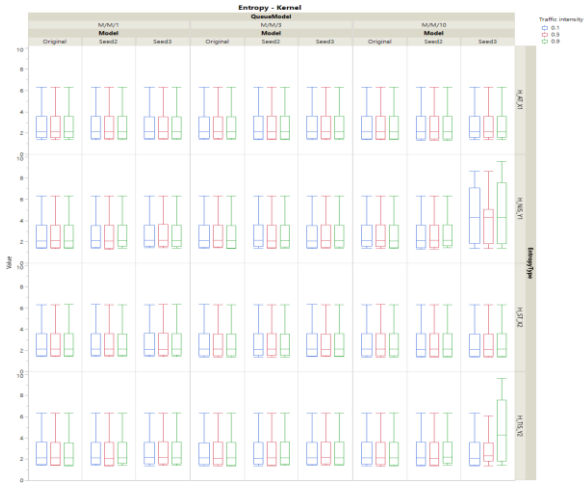

Figure 43. Entropy and MI measures per queue model per traffic-intensity per seed using kernel method values of bandwidth equal to 0.0001, 0.001, and 0.01.

Figure 44. Entropy and MI measures per queue model per traffic-intensity per seed using kernel method with Silverman bandwidth.

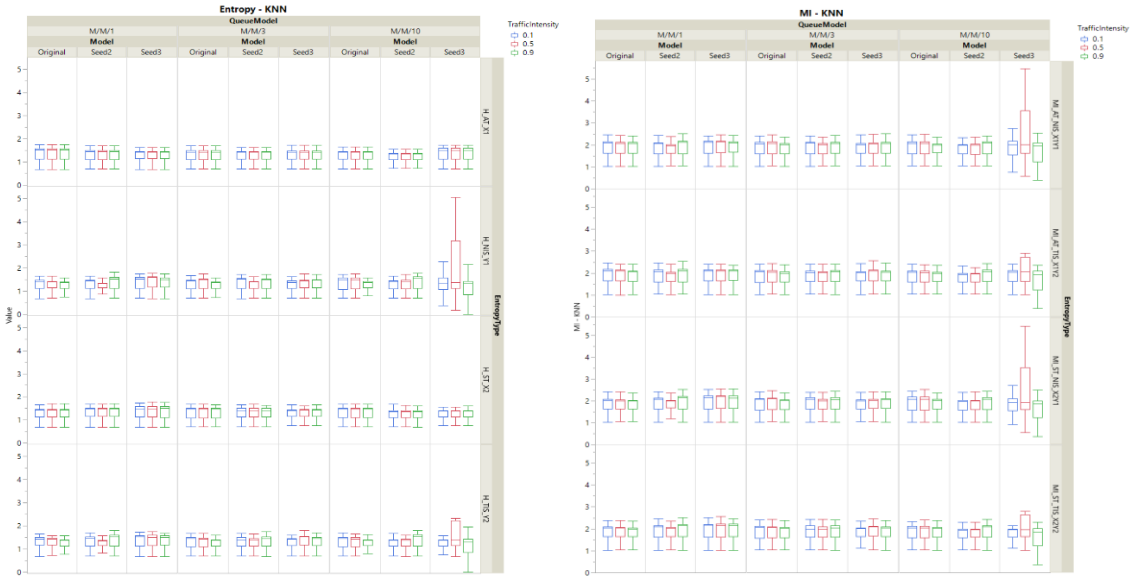

Figure 45. Entropy and MI measures per queue model per traffic-intensity per seed using KNN method with different number of k-nearest neighbors.

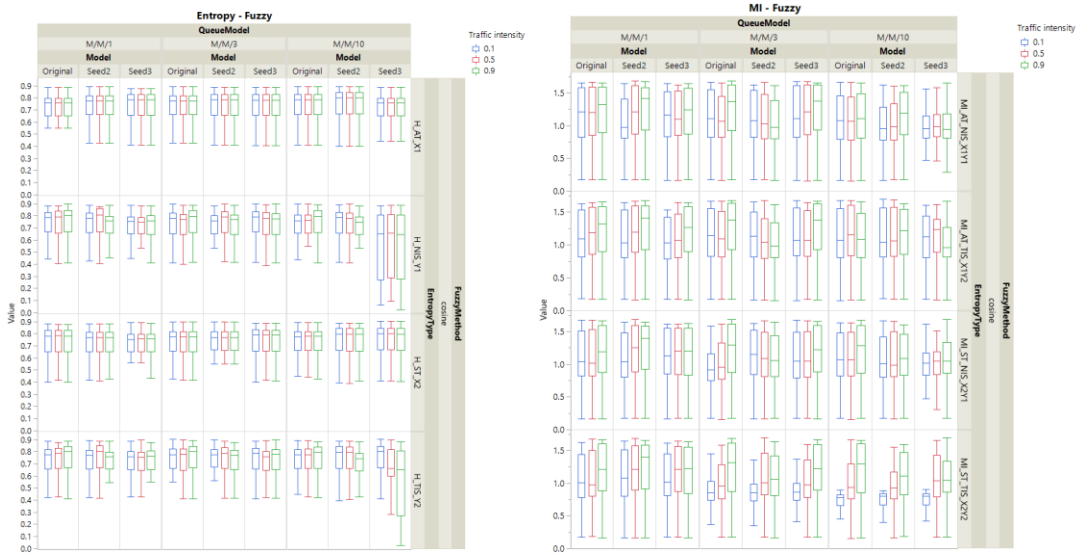

Figure 46. Entropy and MI measures per queue model per traffic-intensity per seed using fuzzy-histogram based method with different values of fuzzy subsets and cosine fuzzy membership function.

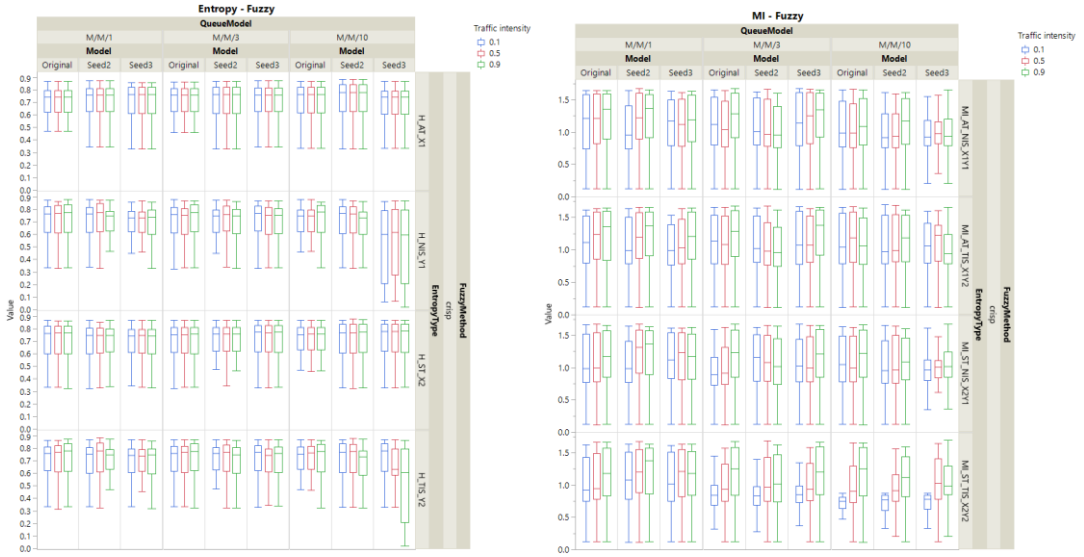

Figure 47. Entropy and MI measures per queue model per traffic-intensity per seed using fuzzy-histogram based method with different values of fuzzy subsets and crisp fuzzy membership function.

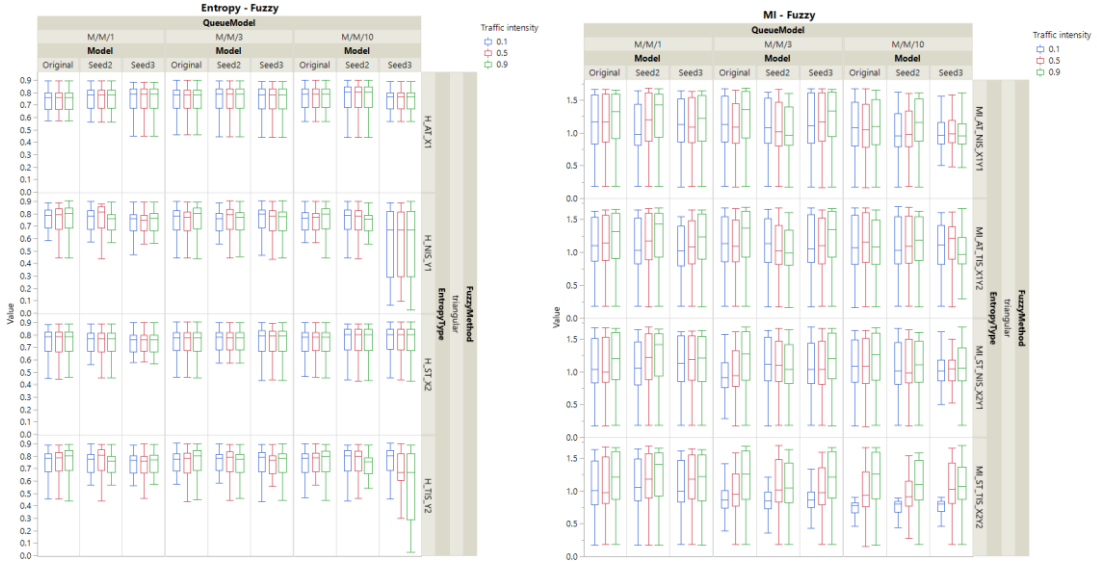

Figure 48. Entropy and MI measures per queue model per traffic-intensity per seed using fuzzy-histogram based method with different values of fuzzy subsets and triangular fuzzy membership function.

Another aspect investigated was how different parameter values would impact the appropriateness of entropy and MI as measures of uncertainty quantification. Experiments #591 to #680 correspond to the initial experiments numbered #1 to #90, but with different parameter values (here, named "number 2", the specific parameter values used in the experiments are given in Table 66). Experiments #711 to #800 correspond to the initial experiments numbered #1 to #90, but with different parameter values (here, named "number 3").

From Figure 49 to Figure 54 one can see that because different values of inputs X_1 and X_2 were used, the entropy of X_1 and X_2 were different among the different experiments: "original", "number 2", and "number 3". However, as expected for a fixed seed, the entropy of X_1 was still equal among the different traffic intensities within each group of experiments. Similar to the observations made when using the histogram-based method, from Figure 49 to Figure 54 one can see that regardless of the method being used to calculate the entropy or MI measures, the traffic intensity does not appear to have a clear relation to the uncertainty of the outputs, as the uncertainty either increases or decreases based on the queue model and changes in the system configurations led to different values of uncertainty without a clear pattern.

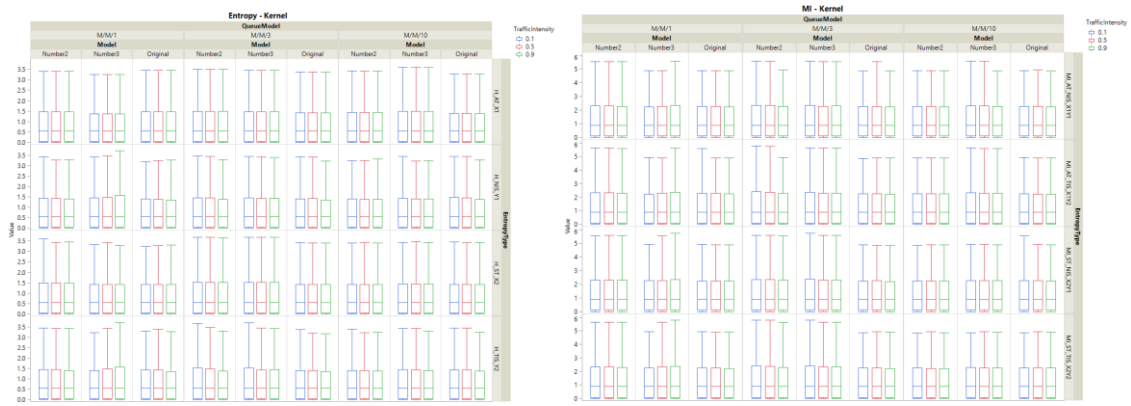

Figure 49. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using kernel method with different values of bandwidth.

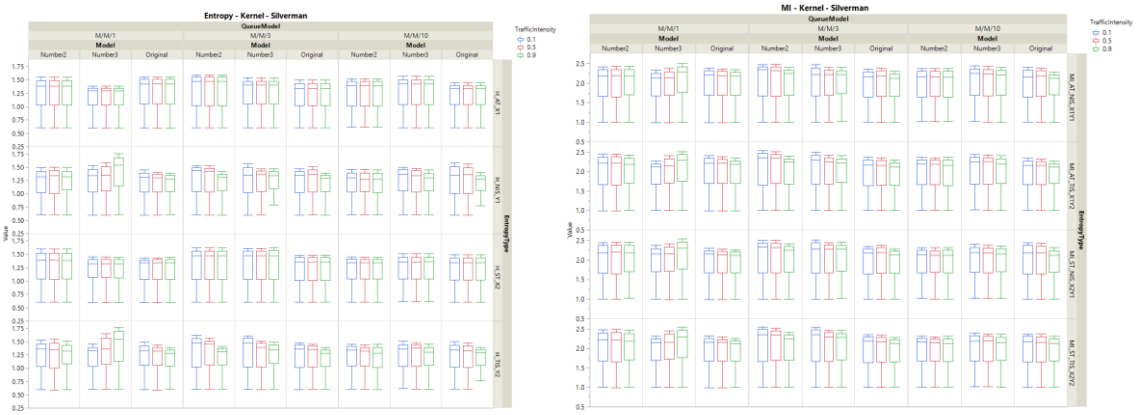

Figure 50. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using kernel method with Silverman bandwidth.

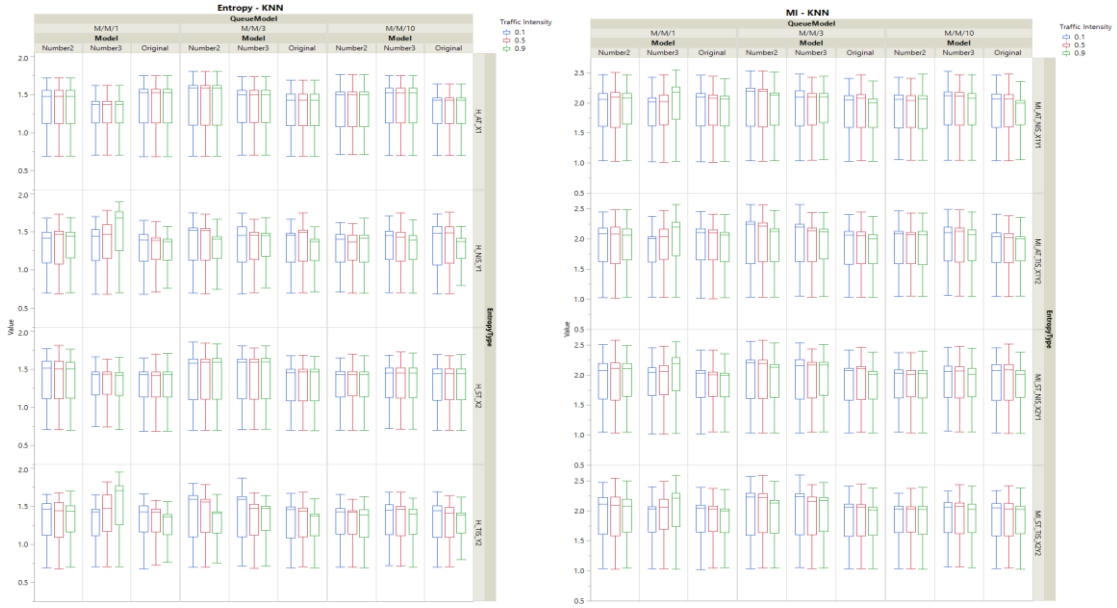

Figure 51. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using KNN method with different values of k-nearest neighbors.

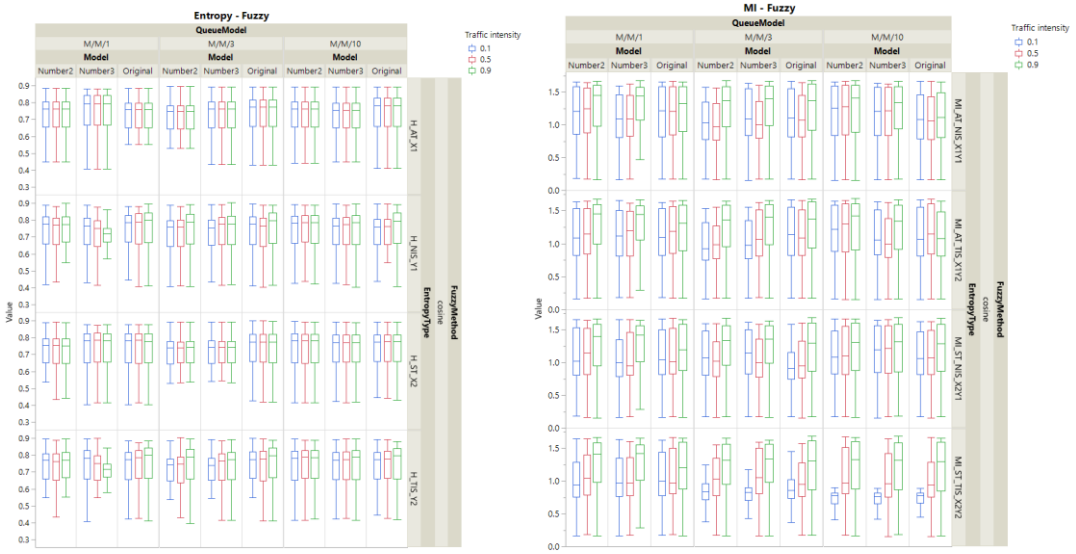

Figure 52. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using fuzzy-histogram based method with different number of fuzzy subsets and cosine membership function.

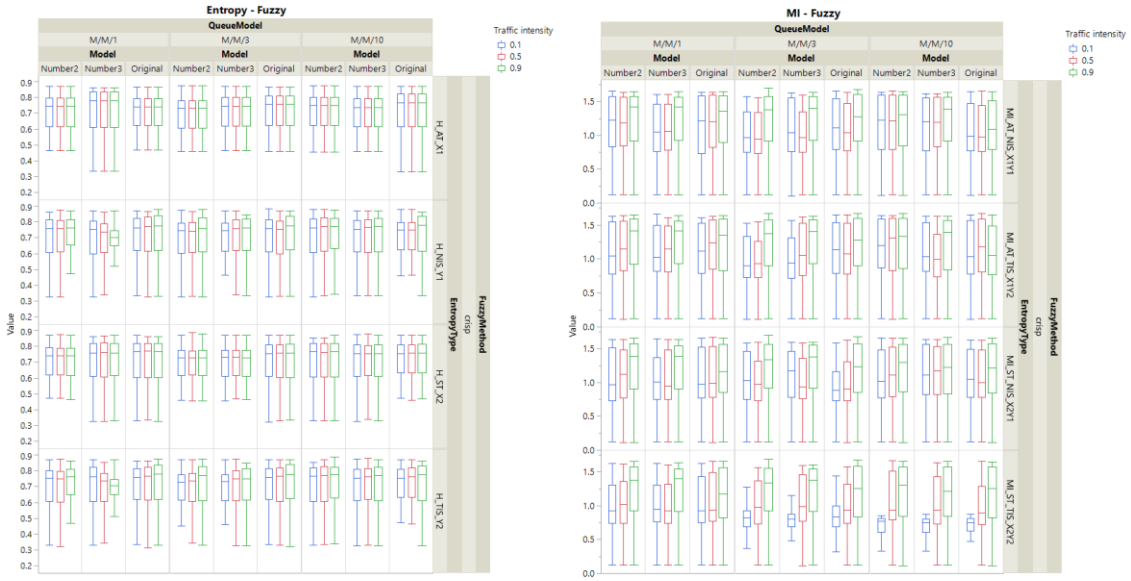

Figure 53. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using fuzzy-histogram based method with different number of fuzzy subsets and crisp membership function.

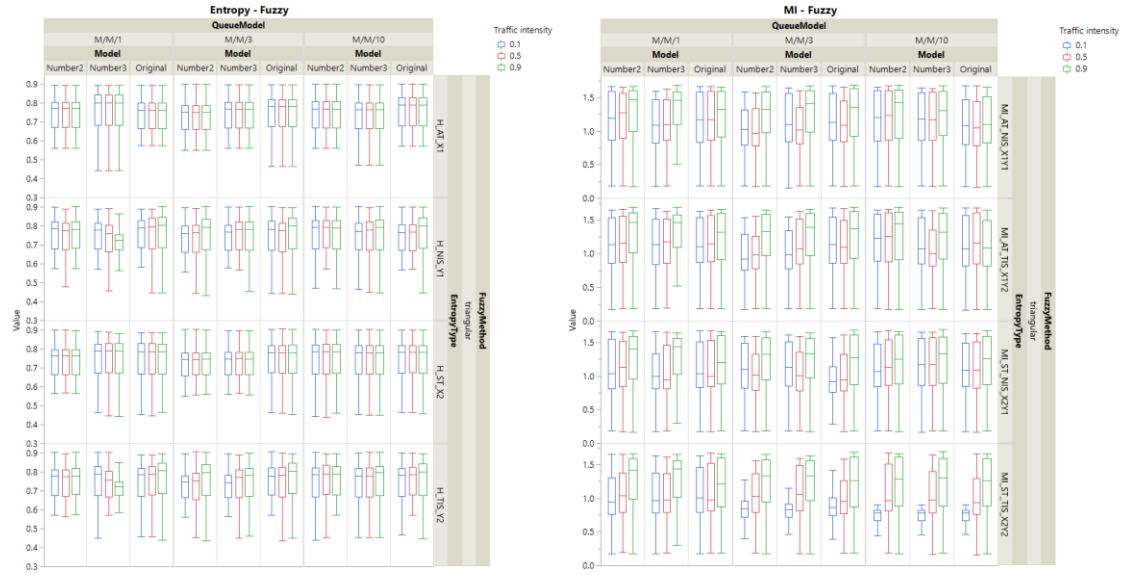

Figure 54. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using fuzzy-histogram based method with different number of fuzzy subsets and triangular membership function.

Finally, the last aspect used to investigate the appropriateness of entropy and MI as measures of uncertainty quantification was whether or not the measures would be able to capture the uncertainty of different systems. For that, two last groups of experiments were run: a CONWIP system and the addition of a third input parameter, namely travel time.

A total of 100 CONWIP experiments were run. Experiments #831 to #930 in Table 66 correspond to the initial experiments numbered #1 to #100, with same service time but constant work in progress. CONWIP systems are systems where the number of items, here the number of customers, is kept constant. Because the number of customers is kept constant, the CONWIP system is expected to have no uncertainty regarding $\widehat{Y_1}$ and, hence, the inputs X_1 and X_2 should have no impact on $\widehat{Y_1}$. Moreover, the customer's time in system is determined by how long it takes to be serviced plus time spent in the queue, if any. The goal of using the CONWIP system is to investigate the effectiveness of the entropy measures in capturing these known characteristics. Here, different from the previous experiments, the arrival process is determined by the service process.

Based on the aforementioned characteristics, one should expect the entropy of the average number of entities in the system, $\widehat{Y_1}$, to be zero. One could expect the entropy of the arrival process, X_1, the service time, X_2, and the average time in the system, $\widehat{Y_2}$ to be equal, as the arrival process and the time in system are dictated by the service time. However, in a simulation model two events, e.g., an arrival and service, cannot occur exactly at the same time. Moreover, the output uncertainty is not only comprised by the input uncertainty but also by some other uncertainties of the system (for instance, the

computational limitation), and, hence, some small differences between X_2 and $\widehat{Y_2}$ are also expected.

From Figure 55, one can observe that when using the kernel method with different values of bandwidth the entropy of $\widehat{Y_1}$ is constant over the different number of replications and it approximates 0 for larger values of bandwidth (bandwidth = 100 or 1,000). However, the entropy of X_1, X_2, and $\widehat{Y_2}$ are also constant and approximate 0 for larger values of bandwidth. The reason for the entropy to approximate 0 has already been discussed in section 3.4.3 and can be an indication of oversmoothing. It is important to discuss why the entropy of $\widehat{Y_1}$ is constant even for lower values of bandwidth. The reason is that in the kernel method the probability density is estimated by $\hat{f}(y_j) = \frac{1}{nh}\sum_{i=1}^{n} K(v)$ where $v = \frac{y_j - y_i}{h}$. In the CONWIP system, $y\ \forall\ i$ are equal, which means $v = 0$, $K(v) \forall i$ are equal, and because n and h are the same $\forall j$, $\hat{f}(y_j)$ are equal $\forall\ j$. Because entropy is approximated using $\widehat{H}(X) = -\frac{1}{n}\sum_{j=1}^{n} \log \hat{f}(X_j)$, regarding of the number of replications, $\widehat{H}(X)$ will be the same for the same value of bandwidth.

One would expect that the inputs X_1 and X_2 should have no impact on $\widehat{Y_1}$ and, hence, $I(X_1; Y_1)$ and $I(X_2; Y_1)$ should be equal to 0. However, from Figure 55 one can see that the MI is not 0 and it slightly varies over the different number of replications for lower values of bandwidth. The reason for this is that MI can be calculated by $I(X_1; Y_1) = H(X_1) + H(Y_1) - H(X_1, Y_1)$. As it was just discussed, in the CONWIP system $H(Y_1)$ is constant regardless of the value of the bandwidth. In the kernel method, $\hat{f}^{Kernel}(x_1, y_1) =$

$\frac{1}{nh_{x_1}h_{y_1}}\sum_{i=1}^{n}\left(K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right)K\left(\frac{y_1-y_{1i}}{h_{y_1}}\right)\right)$. Because $K(v) \in [0,1]$, we know that $K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right)K\left(\frac{y_1-y_{1i}}{h_{y_1}}\right) \leq K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right)$, $K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right)K\left(\frac{y_1-y_{1i}}{h_{y_1}}\right) \leq K\left(\frac{y_1-y_{1i}}{h_{y_1}}\right)$, and if $K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right) = 0$, $K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right)K\left(\frac{y_1-y_{1i}}{h_{y_1}}\right) = 0$. For lower values of bandwidth $h_{x_1}h_{y_1} \leq h_{x_1}$ and $h_{x_1}h_{y_1} \leq h_{y_1}$. Therefore, in this case there is a high chance that $H(X_1, Y_1) \neq H(X_1)$. More precisely, because $K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right)K\left(\frac{y_1-y_{1i}}{h_{y_1}}\right) \leq K\left(\frac{x_1-x_{1i}}{h_{x_1}}\right)$, $\hat{f}(x,y)$ tend to be smaller than $\hat{f}(x)$, even though there is also a decrease in the divisor $h_{x_1}h_{y_1}$. The joint entropy is approximated by $\widehat{H}(X_1, Y_1) = -\frac{1}{n}\sum_{j=1}^{n}\log\left[\frac{1}{1+\hat{f}(x,y)}\right]$. Hence, with the decrease in $\hat{f}(x,y)$, $\frac{1}{1+\hat{f}(x,y)}$ increases and the joint entropy decreases. Hence, for lower values of bandwidth, $H(X_1, Y_1)$ tend to be smaller than $H(X_1)$ and $I(X_1; Y_1) = H(X_1) + H(Y_1) - H(X_1, Y_1)$ will not be constant over the number of replications. As the value of the bandwidth increases, all entropy measures tend to be constant over the number of replications and, consequently, the MI will also tend to be constant.

From Figure 56, one can see that, when using lower values of bandwidth, the entropy of X_1, X_2, and $\widehat{Y_2}$ are equal for low number of replications (number of replications = 10 or 20). As the number of replications increases, the differences between the entropy of X_1, X_2, and $\widehat{Y_2}$ also increase. This result is similar to what was observed with the histogram-based method. There appears to be an interaction between the binwidth (or bandwidth) and the number of replications used. When using the kernel method, the differences between X_1, X_2, and $\widehat{Y_2}$ increases with changes in the number of replications,

while when using the histogram-based method with probability density function the difference increases with changes in the number of bins (binwidth). Differences due to number of bins may indicate that the specific bin is not adequate to estimate the probability. On the other hand, differences due to number of replications may indicate that as the data increases there are more uncertainties in the model and the differences between the inputs and output increase. However, one can argue that if this was the case, this should occur for every probability estimation method used to calculate the entropy measures.

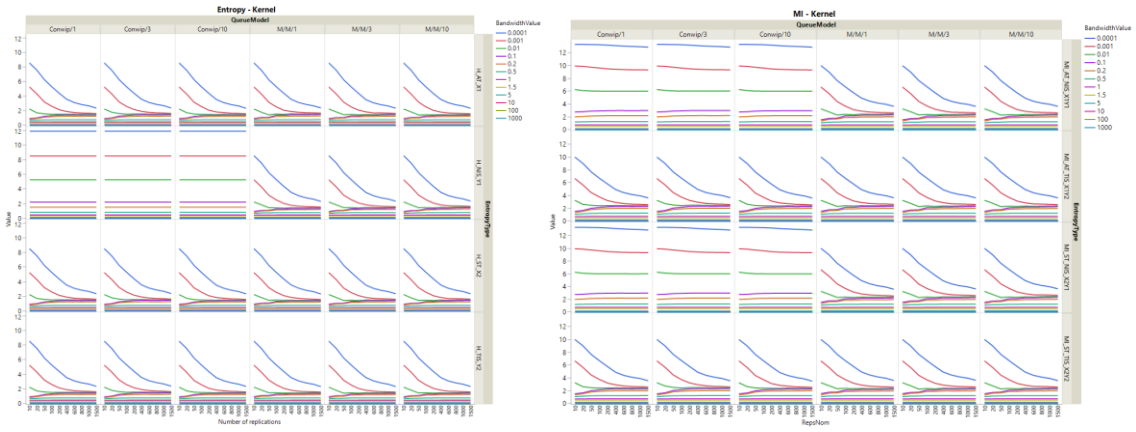

Figure 55. Entropy and MI measures per queue model per number of bins using kernel method with different values of bandwidth (CONWIP vs original experiments).

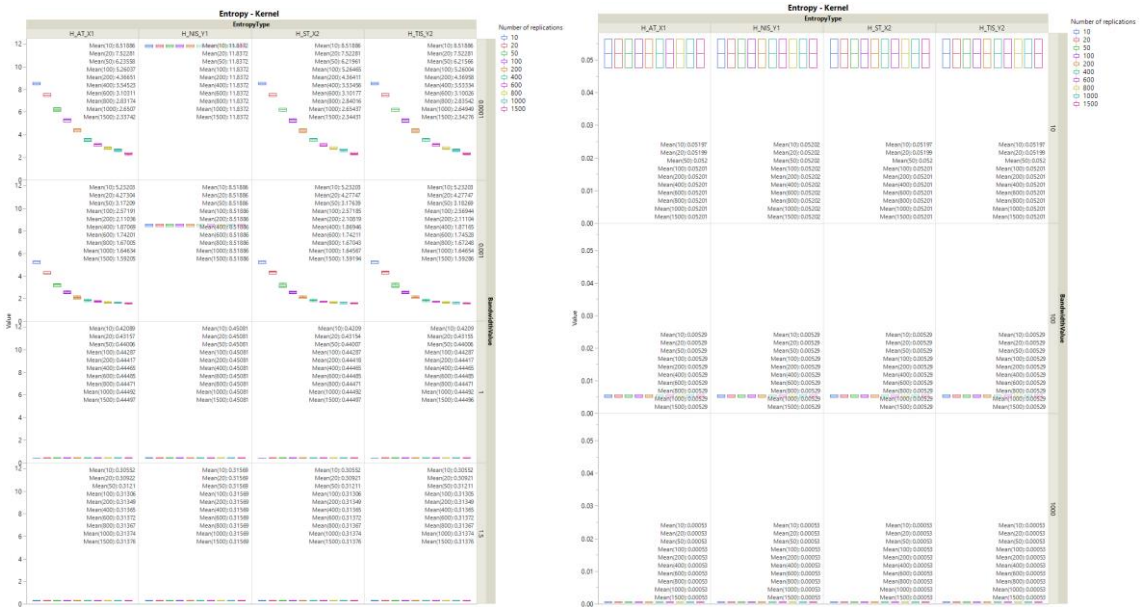

Figure 56. Entropy and MI measures per different values of bandwidth per replication using kernel method for CONWIP systems.

When using the KNN method to calculate the entropy and MI measures, one can observe from Figure 57 that the entropy of $\widehat{Y_1}$ is equal to 0 regardless of the number of k-nearest neighbors used to calculate the measure. Similar behavior is not observed for the entropy of X_1, X_2, and $\widehat{Y_2}$, which is an indication that when using the KNN method the entropy is able to appropriately capture the characteristics of the CONWIP system. However, when analyzing the MI, one can see that although $\mathrm{I}(X_1; Y_1)$ and $\mathrm{I}(X_2; Y_1)$ are closer to 0 than $\mathrm{I}(X_1; Y_2)$ and $\mathrm{I}(X_2; Y_2)$, the $\mathrm{I}(X_1; Y_1)$ and $\mathrm{I}(X_2; Y_1)$ are not equal to 0 as expected in the CONWIP system and, hence, the MI is not appropriately representing the average reduction in uncertainty about the value of X provided by the value of Y. The reason for the MI to not be able to appropriately represent the average reduction in uncertainty comes from how the method is estimated in the multivariate case. For

calculating the entropy, the probability density was estimated using $\hat{f}^{KNN}(y) = \frac{k}{n2\sqrt{\left(y_1 - y_{1_k}\right)^2}}$ and for calculating the MI, the probability density was estimated using $\hat{f}^{KNN}(x;y) = \frac{k}{n\pi\sqrt{\left(x_1 - x_{1_k}\right)^2 + \left(y_1 - y_{1_k}\right)^2}}$. In the CONWIP system, when y_1 is the NIS input, $\left(y_1 - y_{1_k}\right)^2$ is equal to 0 because y_1 is always constant. The MI can be calculated by $\mathrm{I}(X_1;Y_1) = \mathrm{H}(X_1) + H(Y_1) - H(X_1,Y_1)$. Because we have, $\mathrm{H}(Y_1) = 0$, $\mathrm{I}(X_1;Y_1) = \mathrm{H}(X_1) - H(X_1,Y_1)$. However, $\frac{k}{n2\sqrt{\left(x_1 - x_{1_k}\right)^2}} > \frac{k}{n\pi\sqrt{\left(x_1 - x_{1_k}\right)^2 + \left(y_1 - y_{1_k}\right)^2}}$ and, hence, $\mathrm{I}(X_1;Y_1) > 0$. Because $\mathrm{H}(Y_1) = 0$, $\mathrm{I}(X_1;Y_1)$ is usually smaller in the CONWIP system than in the other systems. Similar analysis can be done for $\mathrm{I}(X_2;Y_1)$.

When using the KNN method the entropy of X_1, X_2, and $\widehat{Y_2}$ are different regardless of the number of replications and the number of k-nearest neighbors used. Nevertheless, these differences are not large, which could still be explained due to the computational limitation of the simulation model of generating events simultaneously.

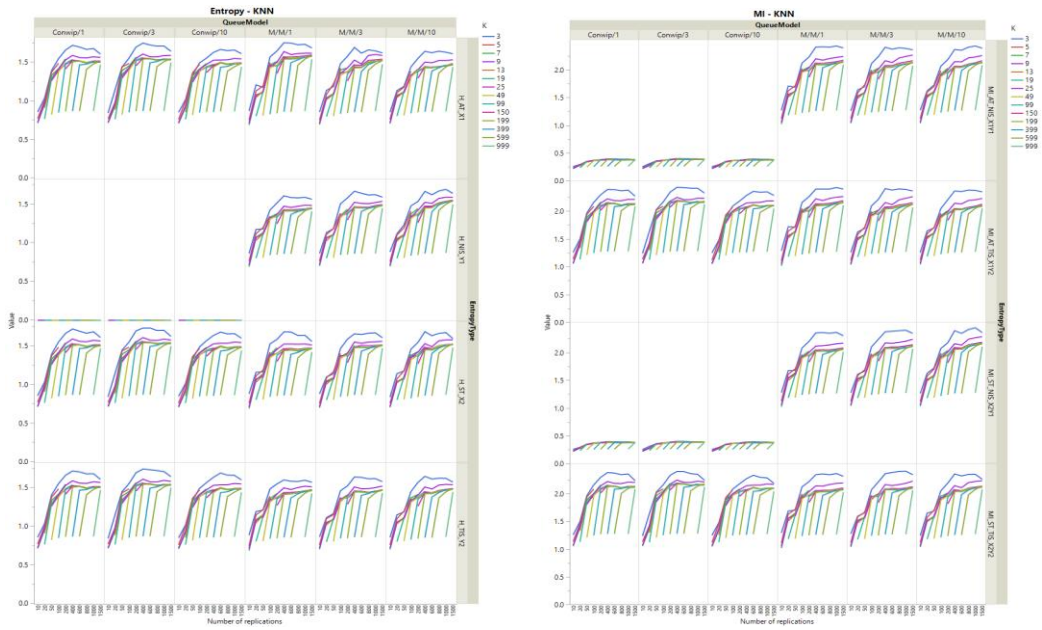

Figure 57. Entropy and MI measures per queue model per number of bins using KNN method with different values k-nearest neighbors (CONWIP vs original experiments).

When using the fuzzy-histogram based method to calculate the entropy and MI measures, one can observe from Figure 58 to Figure 60 that, regardless of the fuzzy membership function considered, the entropy of $\widehat{Y_1}$ is equal to 0. This result is similar to what was observed from the KNN method. The reason for the entropy of $\widehat{Y_1}$ be equal to 0 is due to how the probability density is estimated by the method. The issue here is that because $\widehat{Y_1}$ is constant, the universe is $y = \Omega = a$ instead of $y \in \Omega = [a, b]$, and, consequently, it is not possible to define p fixed nodes of the universe $m_1 < m_2 < \cdots < m_p$. For the universe $y \in \Omega = [a, b]$, $h_k = m_{k+1} - m_k = h = constant$ would be determined by $h = \frac{b-a}{p-1}$. When $\widehat{Y_1}$ is constant, this would lead to $h = 0$. Because the probability density is estimated by $\hat{f}^{fuzzy}(y) = \frac{\sum_{i=1}^{n} \mu_{B_k}(y)}{nh}$, this would result in

indeterminations in the calculation. On the other hand, we know by definition that if $y \notin [m_k, m_{k+1}]$, then $\mu_{B_k}(y) = 0$. Based on that, it was considered that when $\widehat{Y_1}$ is constant, $\mu_{B_k}(y) = 0$. Because $\widehat{H}(Y_1) = -\frac{1}{n}\sum_{j=1}^{n} \log\left[\frac{1}{1+\hat{f}(y)}\right]$, when $\hat{f}(y) = 0$, $\widehat{H}(Y_1) = 0$.

When analyzing the MI, one can see that $\mathrm{I}(X_1; Y_1)$ and $\mathrm{I}(X_2; Y_1)$ are not equal to 0 as expected in the CONWIP system and, hence, the MI is not appropriately representing the average reduction in uncertainty about the value of X provided by the value of Y. The reason again is due to how the probability densities are estimated by the method. The MI can be calculated by $I(X_1; Y_1) = H(X_1) + H(Y_1) - H(X_1, Y_1)$ and $\widehat{H}(X_1, Y_1) = -\frac{1}{n}\sum_{j=1}^{n} \log\left[\frac{1}{1+\hat{f}(x,y)}\right]$. In the fuzzy-histogram based method, $\hat{f}^{fuzzy}(x, y) = \frac{\sum_{i=1}^{n} \mu_{A_k \times B_k}(x,y)}{nh}$, where $\mu_{A_k \times B_k}(x, y) = \min\left(\mu_{A_k}(x), \mu_{B_k}(y)\right)$ is the joint membership function of x and y. Based on that and the fact that $\mu_{B_k}(y) = 0$, we have $\mu_{A_k \times B_k}(x, y) = 0$, $\hat{f}(x_1, y_1) = 0$, and $H(X_1, Y_1) = 0$. Therefore, $I(X_1; Y_1) = H(X_1)$ and not equal to 0.

When using the fuzzy-histogram based method the entropy of X_1, X_2, and $\widehat{Y_2}$ are different regardless of the number of replications and the number of fuzzy subsets used. The exception occurs when crisp membership function is used and 10 replications is considered. In this case, the entropy of X_1, X_2, and $\widehat{Y_2}$ are equal regardless of the number of fuzzy subsets used.

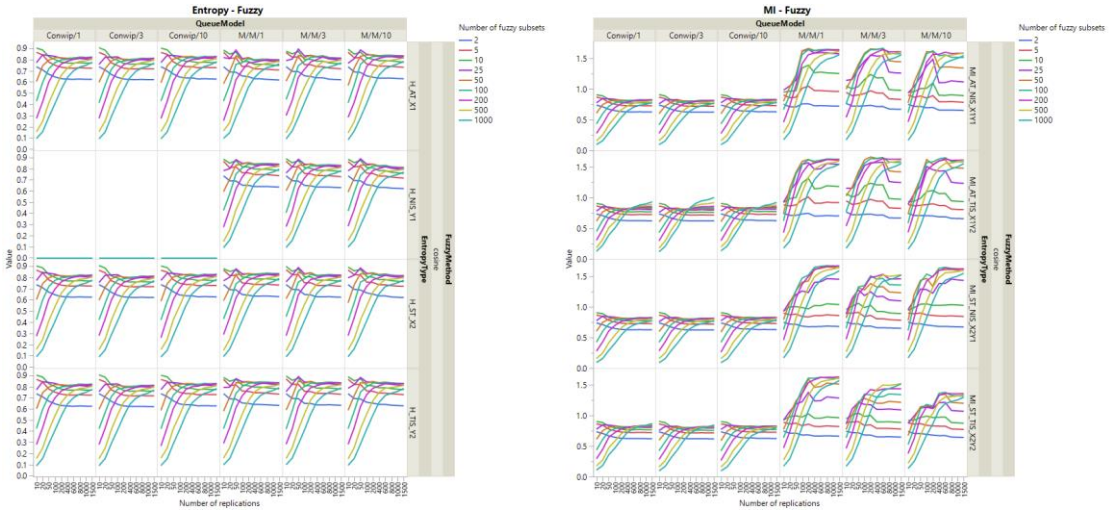

Figure 58. Entropy and MI measures per queue model per number of bins using fuzzy-histogram based method with cosine fuzzy membership function (CONWIP vs original experiments).

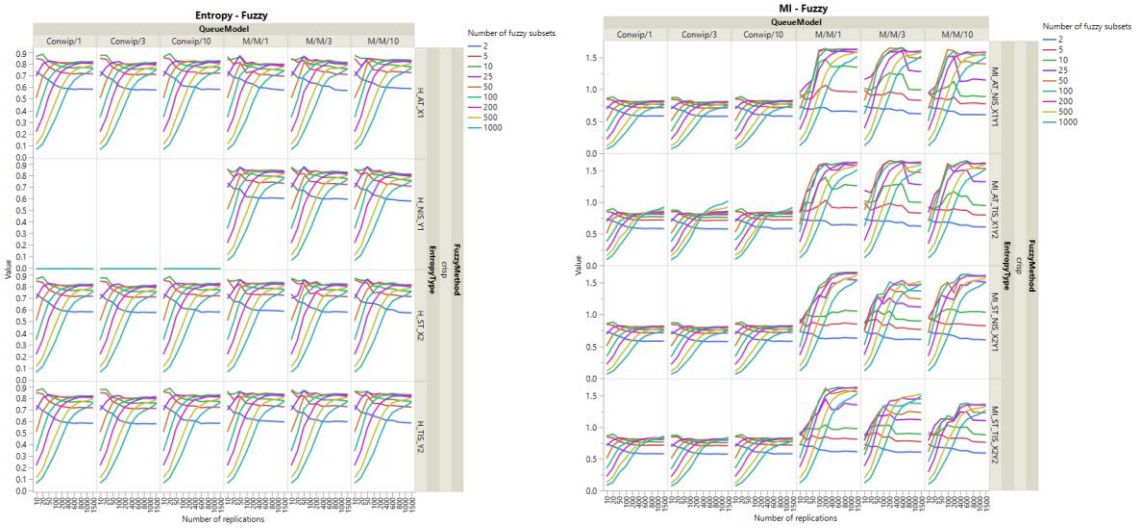

Figure 59. Entropy and MI measures per queue model per number of bins using fuzzy-histogram based method with crisp fuzzy membership function (CONWIP vs original experiments).

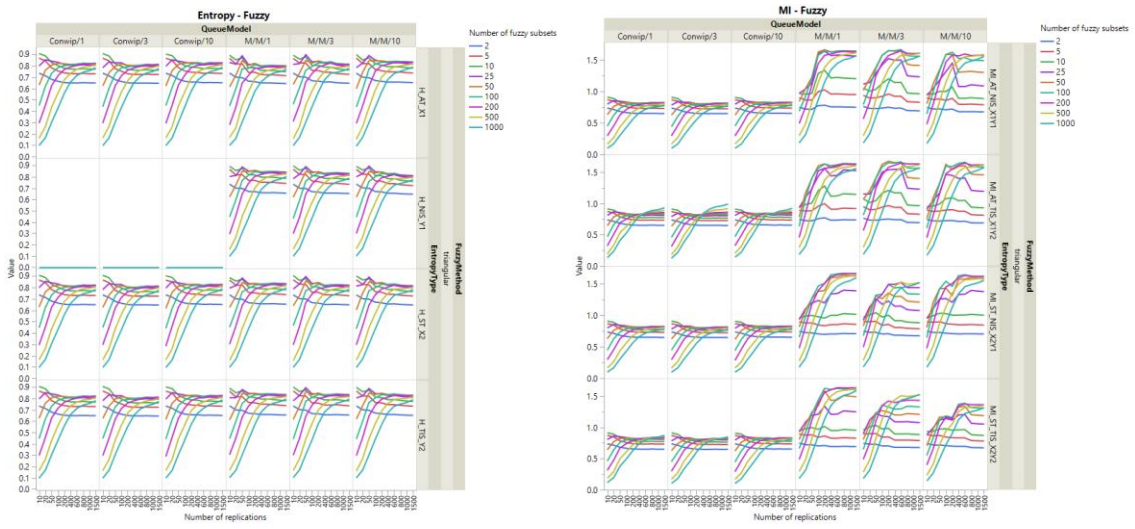

Figure 60. Entropy and MI measures per queue model per number of bins using fuzzy-histogram based method with triangular fuzzy membership function (CONWIP vs original experiments).

In the last group of experiment, a third input parameter X_3 representing the travel time was added to the $M/M/s$ system. A total of 200 experiments with this third input was run: 100 experiments with the added third input as deterministic travel time of 10 minutes and 100 experiments with the added third input as stochastic travel time exponentially distributed with a mean of 10 minutes and fixed seed.

When using the deterministic travel time input, the entropy of X_3 should be zero, as there is no uncertainty associated with the input. Similarly, the input X_3 should bring no reduction in the average uncertainty of the simulation outputs $\widehat{Y_1}$ and $\widehat{Y_2}$, which means that $\mathrm{I}(X_3; Y_1)$ and $\mathrm{I}(X_3; Y_2)$ should be equal to 0. For the stochastic case, there is uncertainty associated with the travel time input and the entropy and MI should capture it.

The observations that can be made when adding the third input deterministic are identical to the ones made from the CONWIP system, as can be seen from Figure 61 to Figure 65.

When using the kernel method with different values of bandwidth the entropy of X_3 is constant over the different number of replications and it approximates 0 for larger values of bandwidth (bandwidth = 1,000) but it is not equal to 0. With respect to the MI, as it was discussed in the CONWIP system, the MI will not be constant over the number of replications because $H(X_3, Y_1)$ tends to be smaller than $H(Y_1)$ and $I(X_3; Y_1) = H(X_3) + H(Y_1) - H(X_3, Y_1)$. The same explanation can be applied to $I(X_3; Y_2)$. Both $I(X_3; Y_1)$ and $I(X_3; Y_2)$ will tend to become constant over the different number of replications and to approximate 0 as the value of the bandwidth becomes large (bandwidth = 1,000). These results of the entropy and MI indicate that the measures are somehow capturing the deterministic behavior of the input X_3, but are not appropriately capturing the lack of uncertainty of this input nor the lack of impact of this input on the outputs, as can be seen in Figure 61.

When using the KNN method, it is possible to observe from Figure 62 that the entropy of X_3 is equal to 0 regardless of the number of k-nearest neighbors used to calculate the measure, but $I(X_3; Y_1)$ and $I(X_3; Y_2)$ are not equal to 0, although these MI measures are closer to 0 than the other MI measures.

Finally, when using the fuzzy-histogram based method, one can observe from Figure 63 to Figure 65 that the entropy of X_3 is equal to 0 regardless of the fuzzy membership function and/or the number of fuzzy subsets used to calculate the measure,

but $\mathrm{I}(X_3;Y_1)$ and $\mathrm{I}(X_3;Y_2)$ are not equal to 0. The KNN method and fuzzy-histogram based method present similar results in detecting the uncertainty of the deterministic inputs and the impact of this input on the outputs. The difference is that in the KNN method the MI is given by $\mathrm{I}(X_3;Y_1) = \mathrm{H}(Y_1) - H(X_3,Y_1)$, because $H(X_3,Y_1) \neq 0$, while in the fuzzy-histogram based method $H(X_3,Y_1) = 0$ and hence, $\mathrm{I}(X_3;Y_1) = \mathrm{H}(Y_1)$.

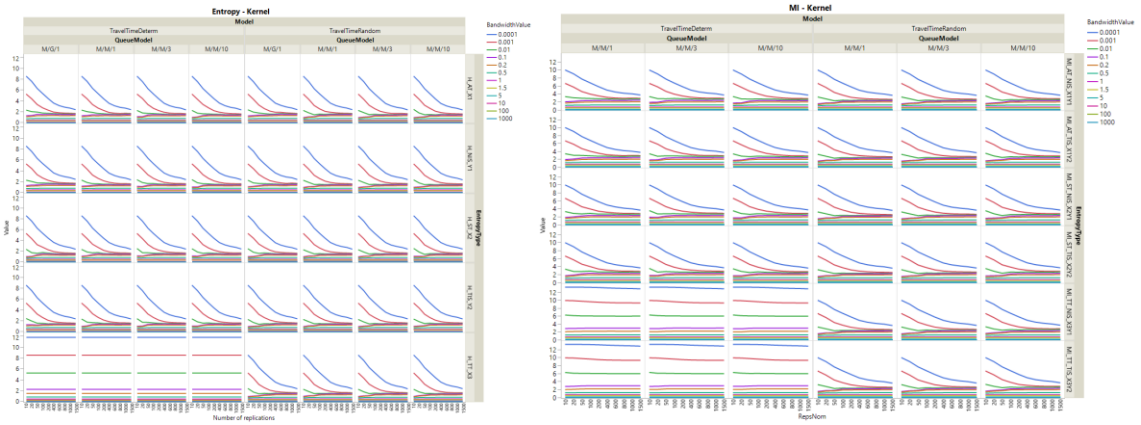

Figure 61. Entropy and MI measures per queue model per number of bins using kernel method with different values of bandwidth (deterministic travel time vs stochastic travel time).

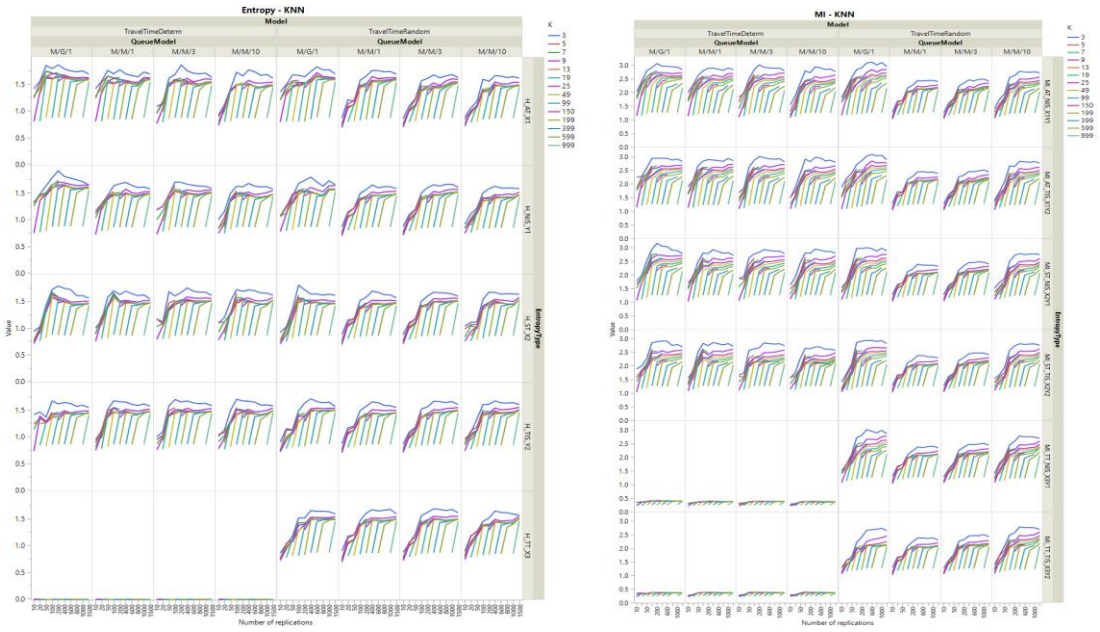

Figure 62. Entropy and MI measures per queue model per number of bins using KNN method with different values of k-nearest neighbors (deterministic travel time vs stochastic travel time).

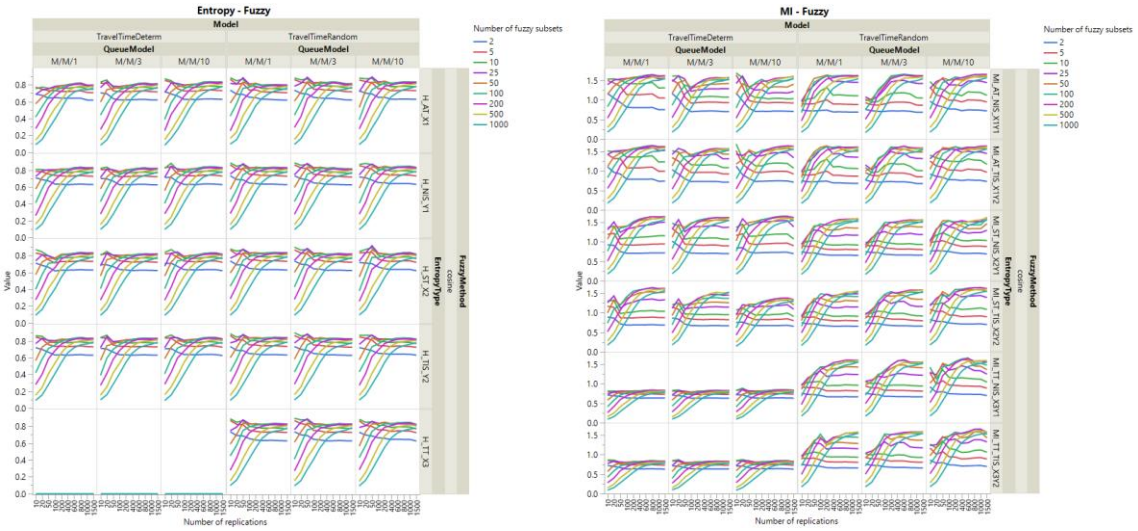

Figure 63. Entropy and MI measures per queue model per number of bins using fuzzy-histogram based method with different number of fuzzy subsets and cosine fuzzy membership function (deterministic travel time vs stochastic travel time).

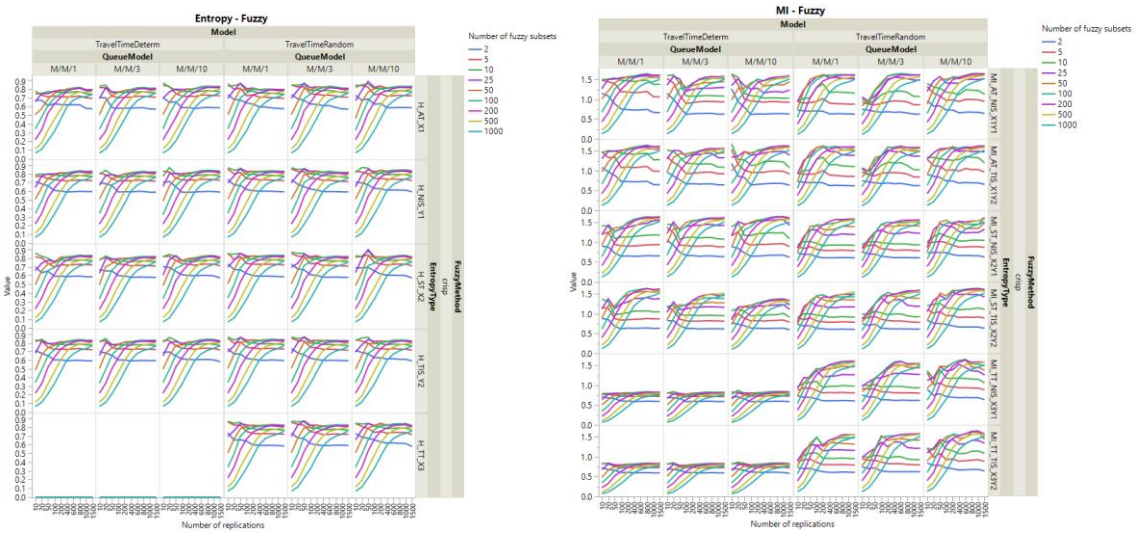

Figure 64. Entropy and MI measures per queue model per number of bins using fuzzy-histogram based method with different number of fuzzy subsets and crisp fuzzy membership function (deterministic travel time vs stochastic travel time).

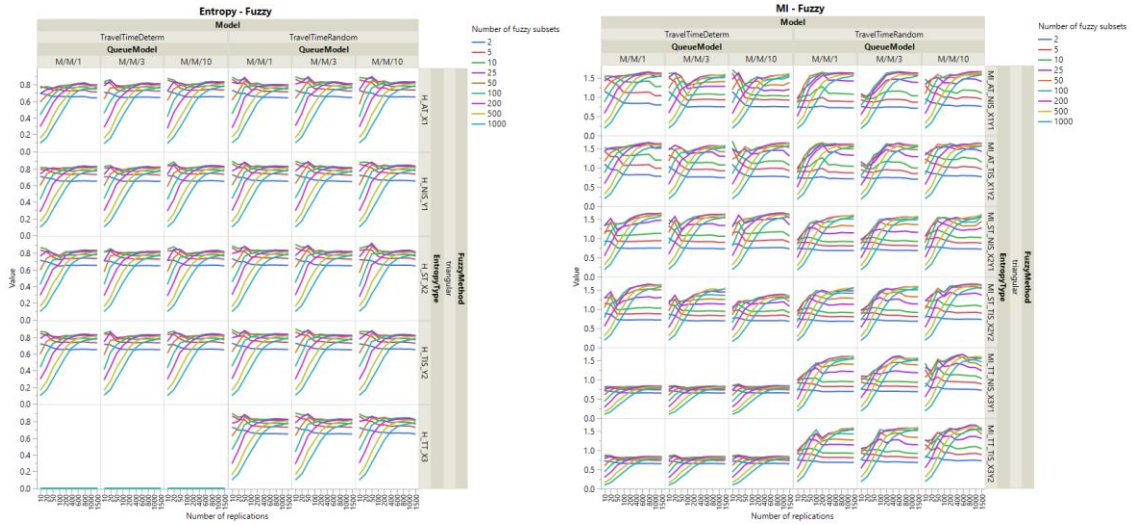

Figure 65. Entropy and MI measures per queue model per number of bins using fuzzy-histogram based method with different number of fuzzy subsets and triangular fuzzy membership function (deterministic travel time vs stochastic travel time).

3.4.7. The histogram-based method

In section 2, the entropy and MI measures were calculated using the histogram-based method. In that section, the use of a data normalization method was proposed to overcome the challenges encountered when entropy is applied to continuous variables. The issue with the data normalization method is that it could only be applied when fixed number of bins was used.

In section 3, a different approach is being proposed based on Jaynes' work. This approach can be applied regardless of method and the number of bins being used. In this section, the impact of different number of bins, different traffic intensities, different seeds, different parameter values, and different systems on the entropy and mutual information measures calculated using the histogram-based method with the approach based on Jaynes' work will be discussed.

From Figure 66 one can see that the entropy and MI measures calculated using the histogram-based method with the approach built on Jaynes' method lead to results similar to the entropy and MI measures calculated using the histogram-based method based on data normalization in terms of behavior of the measures over the number of bins and number of replications. That is, the entropy and MI measures tend to decrease with the increase in the number of bins (or decrease in the binwidth) for the same number of replications and the entropy and MI measures tend to increase with the increase in the number of replications for the same number of bins (or binwidth). The exception is when the number of bins is low (number of bins less than or equal to 25). For low number of bins, the results among the different approaches are somehow different. In the approach based on Jaynes' method, the entropy and MI measures tend to, in fact, decrease with the

increase of the number of replications, while in the data normalization approach the entropy is either constant or increases.

Another big difference among the approaches is that with the approach based on Jaynes' method, the entropy and MI measures are always between 0 and 1, while this is not true for the method based on the data normalization. The reason that entropy is always between 0 and 1 in the approach based on Jaynes' method is that $\hat{f}^{hist}{}_j(x) = \frac{1}{nh}\sum_{i=1}^{n}\mathbf{I}\{x_i \in [t_j, t_{j+1})\} \geq 0$ and, consequently, $0 \leq \frac{1}{1+\hat{f}(x)} \leq 1$, which leads to $0 \leq \widehat{H}(X_1) = -\frac{1}{n}\sum_{j=1}^{n}\log\left[\frac{1}{1+\hat{f}(x)}\right] \leq 1$.

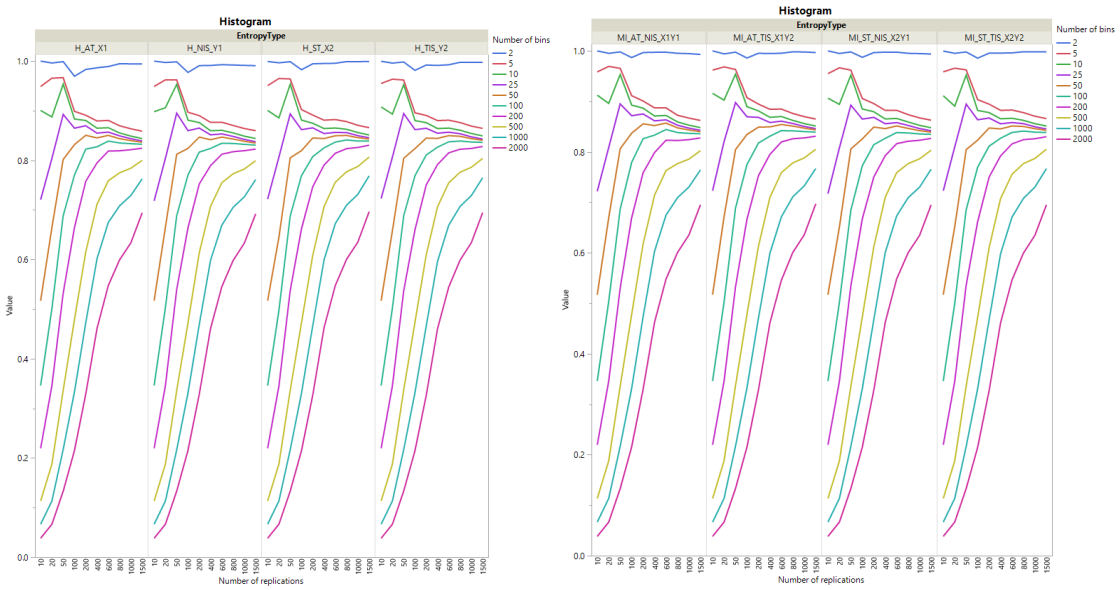

Figure 66. Average of entropy and MI measures per different bandwidths using histogram-based method with different number of bins (experiments #1 to #350).

As it occurred with all the other methods (kernel, KNN, and fuzzy-histogram based) using the approach based on Jaynes' work and also with the histogram-based method with data normalization, the entropy of X_1 was equal among the different traffic intensities when the entropy measure was calculated using the histogram-based method, which indicates that the entropy measure is possibly accurately measuring the information or uncertainty of X_1, as can be seen in Figure 67. Although a fixed seed was also used for the input X_2, it is not correct to expect that X_2 should also have equal entropy among different traffic intensities, as it has already been discussed. As Figure 67 shows, the entropy of X_2 was also able to capture some differences among the different traffic intensities, as expected.

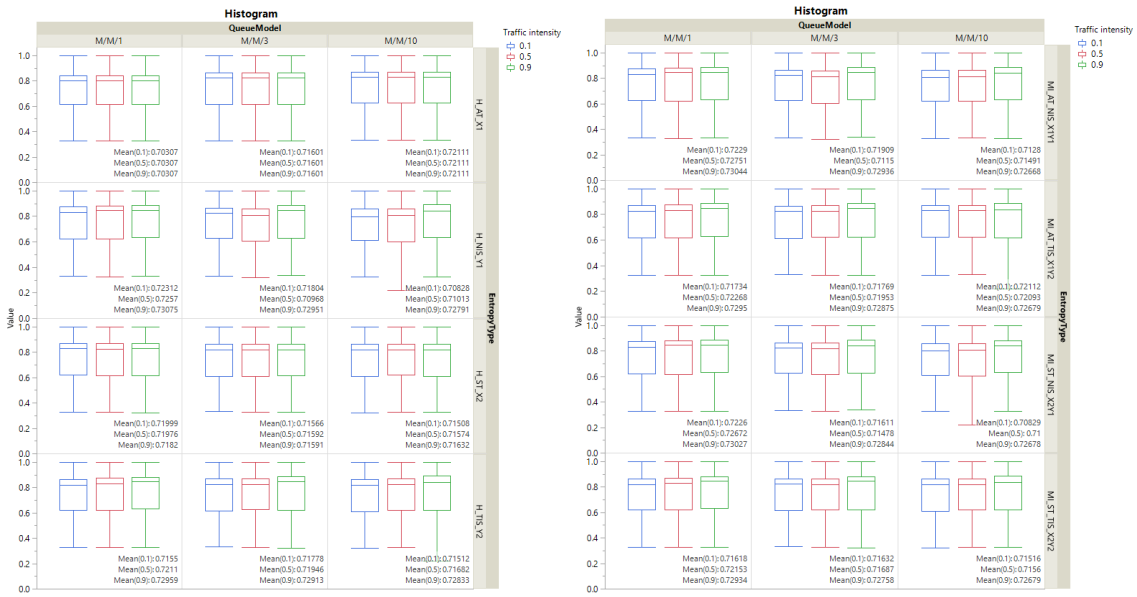

Figure 67. Entropy and MI measures per queue model per traffic-intensity using histogram-based method with different number of bins (experiments #1 to #90).

Figure 68 and Figure 69 show how using different seeds and different parameter values would impact the entropy and MI as measures of uncertainty quantification. As shown in Figure 68, the entropy measures and mutual information vary based on the seed used. For different seeds, there are different uncertainties, as expected. Similar from the results observed with the other methods, when using "seed 3" the entropy of $\hat{Y}_1$ and $\hat{Y}_2$ present the greatest variability among the experiments, which can be potentially used as an indication that the seeds used as "seed 3" are likely not a good combination to be used in the experiments.

From Figure 69 one can see that because different values of inputs X_1 and X_2 were used, the entropy of X_1 and X_2 were different among the different experiments: "original", "number 2", and "number 3". However, as expected for a fixed seed, the entropy of X_1 was still equal among the different traffic intensities within each group of experiments. Similar to the observations made from the other methods, it was not possible to identify a clear relation between the traffic intensity and the uncertainty of the outputs, as the uncertainty either increases or decreases based on the queue model without a clear pattern.

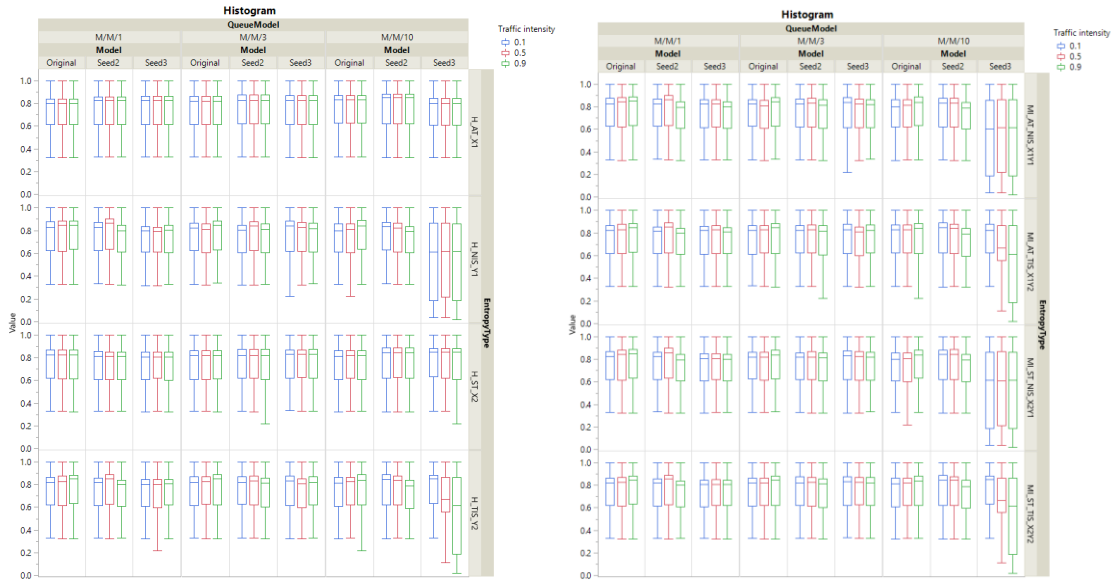

Figure 68. Entropy and MI measures per queue model per traffic-intensity per seed using histogram-based method with different number of bins.

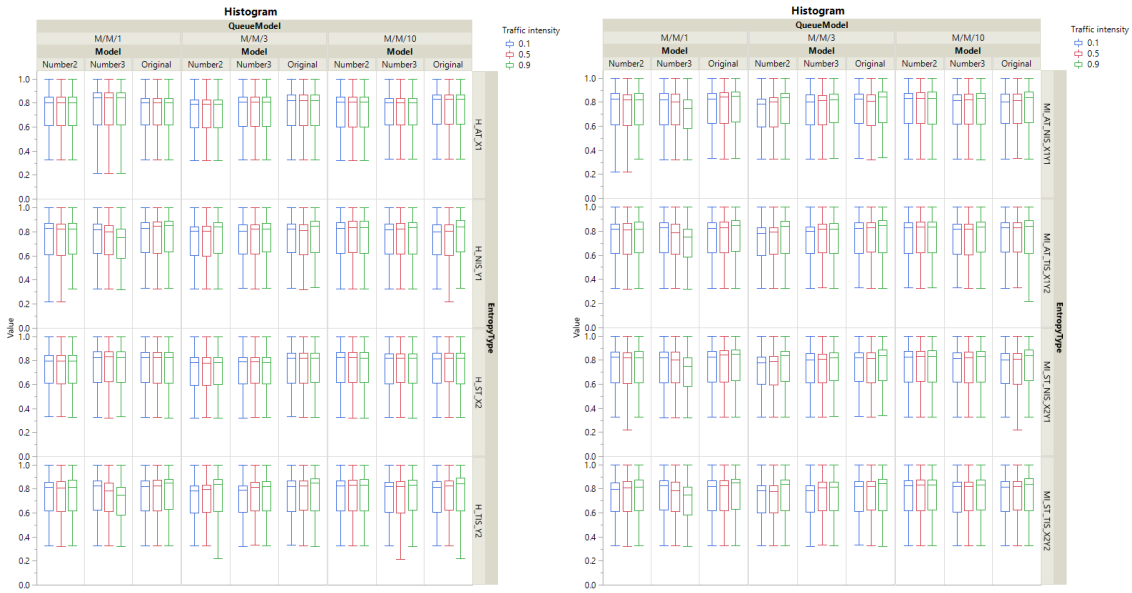

Figure 69. Entropy and MI measures per queue model per traffic-intensity per parameter value experiment using histogram-based method with different number of bins.

The final aspect investigated was whether or not the measures would be able to capture the uncertainty of different systems. More specifically, the CONWIP system and the addition of a third input parameter, namely travel time. Again, the results obtained were very similar to the ones obtained with the histogram-based method with probability density function, fixed number of bins, and data normalization. As can be seen in Figure 70 and Figure 71, when the input or output is deterministic, the entropy is constant over the number of replications and it approximates 0 as the number of bins increases. The MI presents identical behavior to the entropy measures for reasons already discussed. This indicates that similar to the kernel, KNN, and fuzzy-histogram method, the measures calculated using the histogram-based method are somehow capturing the deterministic behavior of the output Y_1 (or input X_3), but are not appropriately capturing the lack of uncertainty of this output (or input) nor the lack of impact of this output (or input) on the inputs (or outputs), as can be seen in Figure 70 and Figure 71.

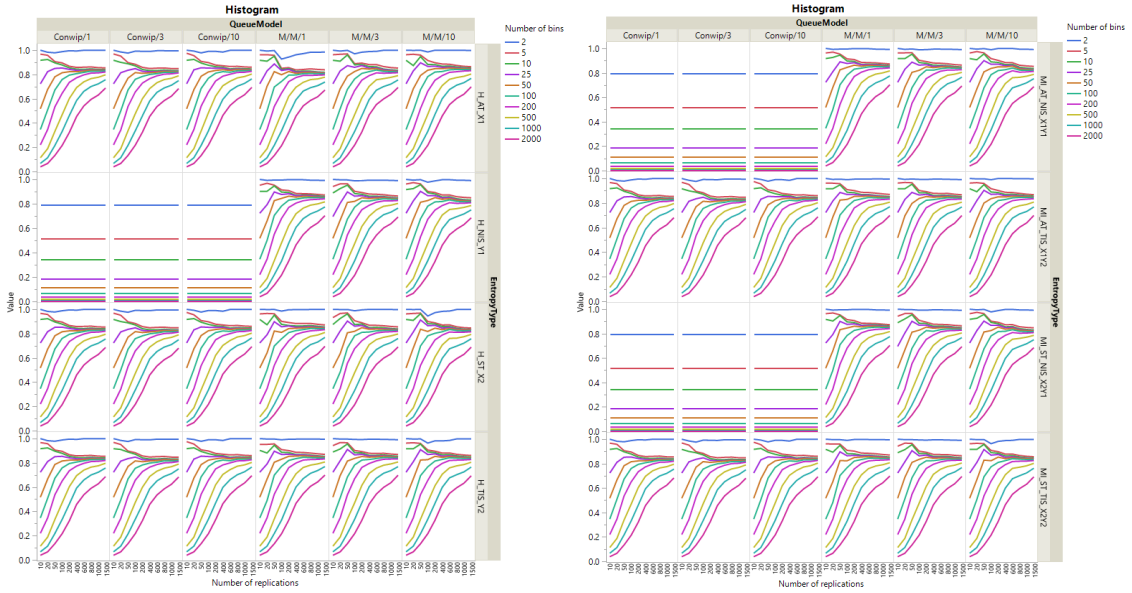

Figure 70. Entropy and MI measures per queue model per number of bins using histogram-based method with different number of bins (CONWIP vs original experiments).

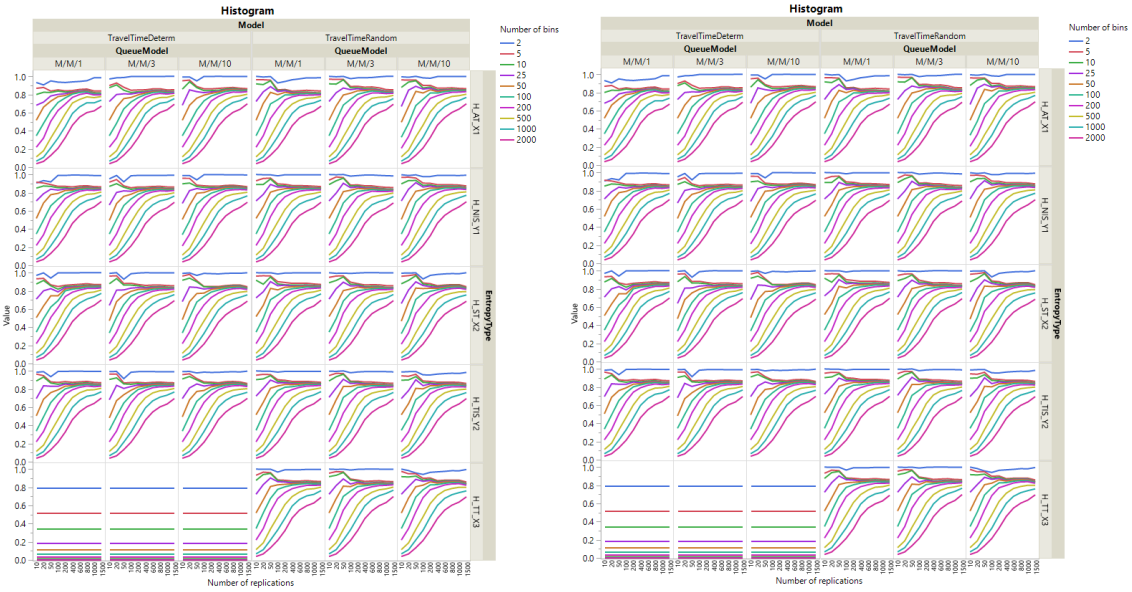

Figure 71. Entropy and MI measures per queue model per number of bins using histogram-based method with different number of bins (deterministic travel time vs stochastic travel time).

3.4.8. Analysis of entropy and MI as a measure of uncertainty quantification in simulation models

In order to assess the potential of entropy and MI as measures of uncertainty quantification in simulation model, the same procedure used in section 2.4.4 was adopted here, where the measures were calculated using the histogram-based method. The procedure consists of comparing the entropy and MI measures results against results of methods commonly applied in the scientific community.

For the entropy measures, the following comparisons were performed:

- The entropy measure (or average entropy measure) detects an increase or decrease in uncertainty with the increase in the number of replications that is in agreement with the detection by the error method being compared to.
- The entropy measure (or average entropy measure) detects the experiment that leads to the maximum uncertainty in agreement with the detection by the error method being compared to.

For the comparisons, four error methods were considered: SAE, SSE, MAE, and MSE. For consistency in the comparison, the average entropy measure was considered instead of entropy measure when MAE or MSE was the error comparison method. This led to a total of eight comparisons per input or output per method used to calculate the entropy measures.

For details about the reasons for the above choice, please refer to section 2.4.4. JMP® and Tableau® software were used to perform the comparisons and analysis. Each comparison was performed for the entropy measures calculated using: (1) the kernel

method with different number of bandwidths; (2) the kernel method with Silverman bandwidth; (3) the KNN method; (4) the fuzzy-histogram based method with cosine fuzzy membership function; (5) the fuzzy-histogram based method with crisp fuzzy membership function; and, (6) the fuzzy-histogram based method with triangular fuzzy membership function.

For the MI measure, the MI results were compared to the results of three other measures of dependence between variables: distance correlation, Pearson correlation, and ${R^2}_{adj}$. To assess the results, the following analyses were performed:

- The MI is capable of identifying the input X_i that has the greatest impact on the uncertainty of the output Y_j in agreement with the measure of dependence.
- The MI is capable of identifying the input X_i that has the least impact on the uncertainty of the output Y_j in agreement with the measure of dependence.

where $i = 1,2$ or $i = 1,2,3$ depending on the scenario being evaluated and $j = 1,2$.

To compare the MI with the measures of dependence an analysis procedure had to be developed because there could be experiments where more than one input would be identified as the one with the greatest impact on the output by either the MI or the other measures of dependence. Similarly, there could be experiments where more than one input would be identified as the one with the least impact on the output by the MI or the other measures of dependence. In these cases, where more than one input would be identified, a procedure to identify whether the MI and the measure of dependence were in agreement was important. This procedure was also described in detailed in section 2.4.4, so it is skipped here.

Table 33 to Table 41 show the results of the comparison between the entropy measures and the SAE, SSE, MAE, and MSE for each of the density estimation methods considered in this section of the dissertation (kernel method, KNN method, and fuzzy-histogram based method).

As shown in Table 33, considering the comparison to detect an increase or decrease in uncertainty the entropy measures calculated using kernel method with different values of bandwidth and different kernel functions do not appear to have results in agreement with the SAE or SSE method for low values of bandwidth (considering the values of the bandwidth experimented – i.e., 0.0001, 0.001, and 0.001). This implies, for instance, that only 2.39% of the time the results of $H(X_1)$ matched the results of the SAE or SSE method for a value of bandwidth of 0.001 when using either the normal or the Epanechnikov kernel functions. The agreement between the entropy measures and the SAE or SSE methods increases as the bandwidth value increases until it reaches its maximum. It is worth noting that the comparisons performed between the entropy measures and the SAE method led to exactly the same results as of the comparisons performed between the entropy measures and the SSE method.

The average of the entropy measures calculated using the kernel method with different values of bandwidth and different kernel functions was compared to MAE and MSE. In this case, it is observed that the performance of the method is, in general, constant over the different values of bandwidth or for the outputs it shows slightly improvement in performance with the increase of the value of the bandwidth. Although the entropy measures when compared against the MAE or MSE perform poorer than when compared

against SAE or SSE, the method can still achieve a performance of about 45% to 55%, with the exception of when considering $H(X_3)$.

Similar observations can be made in terms of the entropy measures performance to detect the experiment that leads to the maximum uncertainty. As shown in Table 34, the ability of the entropy measures in comparison to the SAE and SSE method increases with the increase of the value of the bandwidth until it reaches its maximum and the performance of the average entropy measures when compared to MAE or MSE are constant over the different values of bandwidth. Analogous to what was observed from the results of the entropy measures calculated using the histogram-based method with fixed number of bins and probability density function, the performance of the entropy measures to detect the experiment that leads to the maximum uncertainty is poor when compared to MAE or MSE.

By observing all the results, one can see that a potential good choice for the bandwidth for the queue system under investigation would be a value between 0.2 and 0.5. These values of bandwidth led to close to the best performance, if not the best performance, and could also minimize the risks of oversmoothing that one may encounter when choosing a large value of bandwidth such as 1,000 or even 5.

The results obtained with the kernel method are very similar to the ones obtained with the histogram-based method using fixed number of bins and probability density function in terms of behavior. The difference is that, overall, the kernel method appears to have worse performance when compared to the SAE, SSE, MAE, and MSE than the histogram-based method using fixed number of bins and probability density function.

Another difference is that the performance of the entropy measure calculated using the kernel method when compared with the MAE or MSE is in general constant over the different values of bandwidth, while the performance of the entropy measure calculated using the histogram-based method and probability density function is only constant for lower number of bins. However, when histogram-based method and discrete empirical distribution is considered, the performance of the entropy measure when compared with the MAE or MSE is also constant over the different number of bins.

Table 33. Results from kernel method using different values of bandwidth and different kernel functions for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Function type	Bandwidth value	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	Epanechnikov	0.0001	0.00%	0.00%	44.30%	54.71%
		0.001	2.39%	2.39%	44.30%	54.71%
		0.01	44.30%	44.30%	44.30%	54.71%
		0.1	79.32%	79.32%	44.30%	54.71%
		0.2	87.34%	87.34%	44.30%	54.71%
		0.5	90.86%	90.86%	44.30%	54.71%
		1	90.86%	90.86%	44.30%	54.71%
		1.5	90.86%	90.86%	44.30%	54.71%
		5	90.86%	90.86%	44.30%	54.71%
		10	90.86%	90.86%	44.30%	54.71%
		100	90.86%	90.86%	44.30%	54.71%
		1000	90.86%	90.86%	44.30%	54.71%
	Normal	0.0001	0.00%	0.00%	44.30%	54.71%
		0.001	2.39%	2.39%	44.30%	54.71%
		0.01	41.49%	41.49%	44.30%	54.71%
		0.1	78.90%	78.90%	44.30%	54.71%
		0.2	84.53%	84.53%	44.30%	54.71%
		0.5	90.44%	90.44%	44.30%	54.71%

		1	90.86%	90.86%	44.30%	54.71%
		1.5	90.86%	90.86%	44.30%	54.71%
		5	90.86%	90.86%	44.30%	54.71%
		10	90.86%	90.86%	44.30%	54.71%
		100	90.86%	90.86%	44.30%	54.71%
		1000	90.86%	90.86%	44.30%	54.71%
$H(X_2)$	Epanechnikov	0.0001	0.00%	0.00%	45.43%	52.18%
		0.001	1.69%	1.69%	45.43%	52.18%
		0.01	46.41%	46.41%	45.43%	52.18%
		0.1	80.03%	80.03%	45.43%	52.18%
		0.2	86.92%	86.92%	45.43%	52.18%
		0.5	88.19%	88.19%	45.43%	52.18%
		1	88.19%	88.19%	45.43%	52.18%
		1.5	88.19%	88.19%	45.43%	52.18%
		5	88.19%	88.19%	45.43%	52.18%
		10	88.19%	88.19%	45.43%	52.18%
		100	88.19%	88.19%	45.43%	52.18%
		1000	88.19%	88.19%	45.43%	52.18%
	Normal	0.0001	0.00%	0.00%	45.43%	52.18%
		0.001	1.13%	1.13%	45.43%	52.18%
		0.01	41.21%	41.21%	45.43%	52.18%
		0.1	79.61%	79.61%	45.43%	52.18%
		0.2	83.83%	83.83%	45.43%	52.18%
		0.5	87.48%	87.48%	45.43%	52.18%
		1	88.05%	88.05%	45.43%	52.18%
		1.5	88.19%	88.19%	45.43%	52.18%
		5	88.19%	88.19%	45.43%	52.18%
		10	88.19%	88.19%	45.43%	52.18%
		100	88.19%	88.19%	45.43%	52.18%
		1000	88.19%	88.19%	45.43%	52.18%
$H(X_3)$	Epanechnikov	0.0001	33.33%	33.33%	19.44%	24.44%
		0.001	12.78%	12.78%	19.44%	24.44%
		0.01	32.78%	32.78%	19.44%	24.44%
		0.1	61.11%	61.11%	19.44%	24.44%
		0.2	61.11%	61.11%	19.44%	24.44%
		0.5	57.22%	57.22%	19.44%	24.44%
		1	56.11%	56.11%	19.44%	24.44%
		1.5	50.56%	50.56%	19.44%	24.44%
		5	61.67%	61.67%	19.44%	24.44%
		10	67.22%	67.22%	19.44%	24.44%

		100	89.44%	89.44%	19.44%	24.44%
		1000	89.44%	89.44%	19.44%	24.44%
	Normal	0.0001	27.78%	27.78%	19.44%	24.44%
		0.001	7.22%	7.22%	19.44%	24.44%
		0.01	37.22%	37.22%	19.44%	24.44%
		0.1	50.00%	50.00%	19.44%	24.44%
		0.2	72.22%	72.22%	19.44%	24.44%
		0.5	56.11%	56.11%	19.44%	24.44%
		1	56.11%	56.11%	19.44%	24.44%
		1.5	56.11%	56.11%	19.44%	24.44%
		5	72.78%	72.78%	19.44%	24.44%
		10	78.33%	78.33%	19.44%	24.44%
		100	89.44%	89.44%	19.44%	24.44%
		1000	89.44%	89.44%	19.44%	24.44%
$H(Y_1)$	Epanechnikov	0.0001	0.98%	0.98%	49.37%	54.85%
		0.001	3.52%	3.52%	49.37%	54.85%
		0.01	45.99%	45.99%	49.37%	54.85%
		0.1	83.40%	83.40%	49.37%	54.85%
		0.2	89.45%	89.45%	48.95%	54.43%
		0.5	90.44%	90.44%	48.95%	54.43%
		1	90.30%	90.30%	48.95%	54.43%
		1.5	90.30%	90.30%	48.95%	54.43%
		5	90.30%	90.30%	48.95%	54.43%
		10	90.30%	90.30%	48.95%	54.43%
		100	90.30%	90.30%	48.95%	54.43%
		1000	90.15%	90.15%	48.95%	54.43%
	Normal	0.0001	0.98%	0.98%	49.37%	54.85%
		0.001	2.95%	2.95%	49.37%	54.85%
		0.01	41.07%	41.07%	49.37%	54.85%
		0.1	82.28%	82.28%	49.37%	54.85%
		0.2	87.48%	87.48%	49.09%	54.57%
		0.5	90.01%	90.01%	48.95%	54.43%
		1	90.30%	90.30%	48.95%	54.43%
		1.5	90.15%	90.15%	48.95%	54.43%
		5	90.30%	90.30%	48.95%	54.43%
		10	90.30%	90.30%	48.95%	54.43%
		100	90.30%	90.30%	48.95%	54.43%
		1000	90.30%	90.30%	48.95%	54.43%
$H(Y_2)$	Epanechnikov	0.0001	0.56%	0.56%	48.24%	53.45%
		0.001	2.53%	2.53%	48.38%	53.59%

		0.01	46.84%	46.84%	48.38%	53.59%
		0.1	83.26%	83.26%	48.38%	53.59%
		0.2	87.62%	87.62%	48.10%	53.31%
		0.5	89.17%	89.17%	48.10%	53.31%
		1	88.75%	88.75%	48.10%	53.31%
		1.5	88.75%	88.75%	48.10%	53.31%
		5	88.75%	88.75%	48.10%	53.31%
		10	88.75%	88.75%	48.10%	53.31%
		100	88.75%	88.75%	48.10%	53.31%
		1000	88.61%	88.61%	48.10%	53.31%
	Normal	0.0001	0.56%	0.56%	48.24%	53.45%
		0.001	1.83%	1.83%	48.38%	53.59%
		0.01	42.76%	42.76%	48.38%	53.59%
		0.1	81.86%	81.86%	48.38%	53.59%
		0.2	87.20%	87.20%	48.10%	53.31%
		0.5	88.33%	88.33%	48.10%	53.31%
		1	88.61%	88.61%	48.10%	53.31%
		1.5	88.75%	88.75%	48.10%	53.31%
		5	88.75%	88.75%	48.10%	53.31%
		10	88.75%	88.75%	48.10%	53.31%
		100	88.75%	88.75%	48.10%	53.31%
		1000	88.75%	88.75%	48.10%	53.31%

Table 34. Results from kernel method using different values of bandwidth and different kernel functions for detecting the experiment that leads to the maximum uncertainty.

Entropy	Function type	Bandwidth value	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	Epanechnikov	0.0001	0.00%	0.00%	10.13%	10.13%
		0.001	0.00%	0.00%	10.13%	10.13%
		0.01	0.00%	0.00%	10.13%	10.13%
		0.1	87.34%	87.34%	10.13%	10.13%
		0.2	87.34%	87.34%	10.13%	10.13%
		0.5	87.34%	87.34%	10.13%	10.13%
		1	87.34%	87.34%	10.13%	10.13%
		1.5	87.34%	87.34%	10.13%	10.13%
		5	87.34%	87.34%	10.13%	10.13%
		10	87.34%	87.34%	10.13%	10.13%

		100	87.34%	87.34%	10.13%	10.13%
		1000	87.34%	87.34%	10.13%	10.13%
		0.0001	0.00%	0.00%	10.13%	10.13%
		0.001	0.00%	0.00%	10.13%	10.13%
		0.01	0.00%	0.00%	10.13%	10.13%
		0.1	87.34%	87.34%	10.13%	10.13%
		0.2	87.34%	87.34%	10.13%	10.13%
	Normal	0.5	87.34%	87.34%	10.13%	10.13%
		1	87.34%	87.34%	10.13%	10.13%
		1.5	87.34%	87.34%	10.13%	10.13%
		5	87.34%	87.34%	10.13%	10.13%
		10	87.34%	87.34%	10.13%	10.13%
		100	87.34%	87.34%	10.13%	10.13%
		1000	87.34%	87.34%	10.13%	10.13%
		0.0001	0.00%	0.00%	2.53%	0.00%
		0.001	0.00%	0.00%	2.53%	0.00%
		0.01	0.00%	0.00%	2.53%	0.00%
		0.1	89.87%	89.87%	2.53%	0.00%
		0.2	89.87%	89.87%	2.53%	0.00%
	Epanechnikov	0.5	92.41%	92.41%	2.53%	0.00%
		1	92.41%	92.41%	2.53%	0.00%
		1.5	92.41%	92.41%	2.53%	0.00%
		5	92.41%	92.41%	2.53%	0.00%
		10	92.41%	92.41%	2.53%	0.00%
		100	92.41%	92.41%	2.53%	0.00%
$H(X_2)$		1000	92.41%	92.41%	2.53%	0.00%
		0.0001	0.00%	0.00%	2.53%	0.00%
		0.001	0.00%	0.00%	2.53%	0.00%
		0.01	0.00%	0.00%	2.53%	0.00%
		0.1	89.87%	89.87%	2.53%	0.00%
		0.2	89.87%	89.87%	2.53%	0.00%
		0.5	89.87%	89.87%	2.53%	0.00%
	Normal	1	91.14%	91.14%	2.53%	0.00%
		1.5	92.41%	92.41%	2.53%	0.00%
		5	92.41%	92.41%	2.53%	0.00%
		10	92.41%	92.41%	2.53%	0.00%
		100	92.41%	92.41%	2.53%	0.00%
		1000	92.41%	92.41%	2.53%	0.00%
		0.0001	0.00%	0.00%	50.00%	50.00%
$H(X_3)$	Epanechnikov	0.001	0.00%	0.00%	50.00%	50.00%

		0.01	0.00%	0.00%	50.00%	50.00%
		0.1	45.00%	45.00%	50.00%	50.00%
		0.2	45.00%	45.00%	50.00%	50.00%
		0.5	45.00%	45.00%	50.00%	50.00%
		1	45.00%	45.00%	50.00%	50.00%
		1.5	45.00%	45.00%	50.00%	50.00%
		5	95.00%	95.00%	50.00%	50.00%
		10	45.00%	45.00%	50.00%	50.00%
		100	95.00%	95.00%	50.00%	50.00%
		1000	95.00%	95.00%	50.00%	50.00%
		0.0001	0.00%	0.00%	50.00%	50.00%
		0.001	0.00%	0.00%	50.00%	50.00%
		0.01	0.00%	0.00%	50.00%	50.00%
		0.1	45.00%	45.00%	50.00%	50.00%
		0.2	95.00%	95.00%	50.00%	50.00%
	Normal	0.5	45.00%	45.00%	50.00%	50.00%
		1	45.00%	45.00%	50.00%	50.00%
		1.5	45.00%	45.00%	50.00%	50.00%
		5	45.00%	45.00%	50.00%	50.00%
		10	45.00%	45.00%	50.00%	50.00%
		100	95.00%	95.00%	50.00%	50.00%
		1000	95.00%	95.00%	50.00%	50.00%
		0.0001	1.27%	1.27%	13.92%	11.39%
		0.001	2.53%	2.53%	13.92%	11.39%
		0.01	3.80%	3.80%	13.92%	11.39%
		0.1	89.87%	89.87%	13.92%	11.39%
		0.2	91.14%	91.14%	13.92%	11.39%
	Epanechnikov	0.5	91.14%	91.14%	13.92%	11.39%
		1	91.14%	91.14%	13.92%	11.39%
		1.5	91.14%	91.14%	13.92%	11.39%
$H(Y_1)$		5	91.14%	91.14%	13.92%	11.39%
		10	91.14%	91.14%	13.92%	11.39%
		100	91.14%	91.14%	13.92%	11.39%
		1000	91.14%	91.14%	13.92%	11.39%
		0.0001	1.27%	1.27%	13.92%	11.39%
		0.001	2.53%	2.53%	13.92%	11.39%
		0.01	3.80%	3.80%	13.92%	11.39%
	Normal	0.1	89.87%	89.87%	13.92%	11.39%
		0.2	91.14%	91.14%	13.92%	11.39%
		0.5	89.87%	89.87%	13.92%	11.39%

		1	91.14%	91.14%	13.92%	11.39%
		1.5	91.14%	91.14%	13.92%	11.39%
		5	91.14%	91.14%	13.92%	11.39%
		10	91.14%	91.14%	13.92%	11.39%
		100	91.14%	91.14%	13.92%	11.39%
		1000	91.14%	91.14%	13.92%	11.39%
$H(Y_2)$	Epanechnikov	0.0001	1.27%	1.27%	12.66%	6.33%
		0.001	1.27%	1.27%	12.66%	6.33%
		0.01	2.53%	2.53%	12.66%	6.33%
		0.1	94.94%	94.94%	12.66%	6.33%
		0.2	96.20%	96.20%	12.66%	6.33%
		0.5	93.67%	93.67%	12.66%	6.33%
		1	93.67%	93.67%	12.66%	6.33%
		1.5	93.67%	93.67%	12.66%	6.33%
		5	93.67%	93.67%	12.66%	6.33%
		10	93.67%	93.67%	12.66%	6.33%
		100	93.67%	93.67%	12.66%	6.33%
		1000	93.67%	93.67%	12.66%	6.33%
	Normal	0.0001	1.27%	1.27%	12.66%	6.33%
		0.001	1.27%	1.27%	12.66%	6.33%
		0.01	1.27%	1.27%	12.66%	6.33%
		0.1	94.94%	94.94%	12.66%	6.33%
		0.2	94.94%	94.94%	12.66%	6.33%
		0.5	94.94%	94.94%	12.66%	6.33%
		1	93.67%	93.67%	12.66%	6.33%
		1.5	93.67%	93.67%	12.66%	6.33%
		5	93.67%	93.67%	12.66%	6.33%
		10	93.67%	93.67%	12.66%	6.33%
		100	93.67%	93.67%	12.66%	6.33%
		1000	93.67%	93.67%	12.66%	6.33%

The entropy measures calculated using the kernel method and Silverman bandwidth showed similar or better results than the entropy measures calculated using the kernel method with different values of bandwidth and different kernel functions for either detecting an increase or decrease in uncertainty with the increase in the number of

replications or for detecting the experiment that leads to the maximum uncertainty, as shown in Table 35 and Table 36, respectively. In general, when the entropy measures calculated using the kernel method and Silverman bandwidth were compared to MAE and MSE, the results were identical to the results when the entropy measures were calculated using the kernel method with different values of bandwidth and different kernel functions. When the comparison was against SAE and SSE, the entropy measures calculated using the kernel method and Silverman bandwidth outperformed the entropy measures calculated using the kernel method with different values of bandwidth when Epanechnikov kernel function is used. The only exception was when travel time input was considered. A possible explanation for the better performance of the kernel method using Silverman bandwidth and Epanechnikov kernel function than the kernel method with different values of bandwidth is that the Silverman's rule of thumb is a choice of the bandwidth that minimizes the mean integrated squared error and the Epanechnikov kernel function leads to the lowest asymptotic MSE.

Table 35. Results from kernel method using Silverman bandwidth and different kernel functions for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Kernel function	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	Epanechnikov	94.80%	94.80%	44.30%	54.71%
	Normal	86.92%	86.92%	44.30%	54.71%
$H(X_2)$	Epanechnikov	93.67%	93.67%	45.43%	52.18%
	Normal	88.19%	88.19%	45.43%	52.18%
$H(X_3)$	Epanechnikov	43.33%	43.33%	19.44%	24.44%

	Normal	40.00%	40.00%	19.44%	24.44%
$H(Y_1)$	Epanechnikov	96.06%	96.06%	49.37%	54.85%
	Normal	90.72%	90.72%	49.37%	54.85%
$H(Y_2)$	Epanechnikov	95.64%	95.64%	48.38%	53.59%
	Normal	89.03%	89.03%	48.38%	53.59%

Table 36. Results from kernel method using Silverman bandwidth and different kernel functions for detecting the experiment that leads to the maximum uncertainty.

Entropy	Function type	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	Epanechnikov	98.73%	98.73%	10.13%	10.13%
	Normal	88.61%	88.61%	10.13%	10.13%
$H(X_2)$	Epanechnikov	100.00%	100.00%	2.53%	0.00%
	Normal	89.87%	89.87%	2.53%	0.00%
$H(X_3)$	Epanechnikov	50.00%	50.00%	50.00%	50.00%
	Normal	45.00%	45.00%	50.00%	50.00%
$H(Y_1)$	Epanechnikov	100.00%	100.00%	13.92%	11.39%
	Normal	91.14%	91.14%	13.92%	11.39%
$H(Y_2)$	Epanechnikov	100.00%	100.00%	12.66%	6.33%
	Normal	96.20%	96.20%	12.66%	6.33%

For the KNN method, in order to assess whether the entropy measures detect an increase or decrease in uncertainty with the increase in the number of replications that is in agreement with the detection by the error method being compared to, the same number of k-nearest neighbors should be used by each experiment, which was not what was initially proposed in section 3.3.2. In section 3.4.4, the goal was to have different values of k that would correspond to low, medium, and high values in comparison with the amount of data available in each experiment to assess the direct impact of k on the entropy

and MI measures. In this section, because the same value of k is necessary, the following values were considered $k = 1,2,3,4,5,6,7,8,9$.

From Table 37, Table 39, Table 40, and Table 41 one can see that both the KNN and the fuzzy-histogram-based methods led to results similar to the kernel method for detecting an increase or decrease in uncertainty with the increase in the number of replications and for detecting the experiment that leads to the maximum uncertainty.

As shown in Table 37, considering the comparison to detect an increase or decrease in uncertainty, the entropy measures calculated using the KNN method with different values of k-nearest neighbors appears to have better results in agreement with the SAE or SSE method for low values of k-nearest neighbors than for low values of bandwidth in the kernel method. The fuzzy-histogram based method also showed slightly better results in agreement with the SAE and SSE method for low values of number of fuzzy subsets than for low values of bandwidth in the kernel method, as one can see in Table 40. The agreement between the entropy measures and the SAE or SSE methods increases as the number of k-nearest neighbors or number of fuzzy subsets increases. Similar to the kernel method, the comparisons performed between the entropy measures and the SAE method led to exactly the same results as of the comparisons performed between the entropy measures and the SSE method when using either the KNN or the fuzzy-histogram based method.

The average of the entropy measures calculated using the KNN and the fuzzy-histogram based methods was compared to MAE and MSE. In this case, it is observed that the performance of the method is, with a few exceptions, constant over the different values

of k-nearest neighbors and fuzzy subsets. Similar to the kernel method, the entropy measures when compared against the MAE or MSE also perform poorer than when compared against SAE or SSE, but the method can still achieve a performance of about 45% to 50%. The only exception is when $H(X_3)$ is considered, in which case the performance is actually about 70%.

Similar observations can be made in terms of the entropy measures performance to detect the experiment that leads to the maximum uncertainty. As shown in Table 39 and Table 41, the ability of the entropy measures in comparison to the SAE and SSE method increases with the increase in the number of k-nearest neighbors or the number of fuzzy subsets, with the exception of $H(X_1)$ when using the KNN method, and the performance of the average entropy measures when compared to MAE or MSE are constant over the different number of k-nearest neighbors or number of fuzzy subsets. Analogous to what was observed from the results of the entropy measures calculated using the histogram-based method with fixed number of bins and probability density function or from the kernel method, the performance of the entropy measures to detect the experiment that leads to the maximum uncertainty is poor when compared to MAE or MSE. More precisely, the performance is between 0 to 14%, with the exception of $H(X_3)$.

By observing all the results, one can see that a potential good choice for the k-nearest neighbors for the queue system under investigation would be a value of 9 and a potential good choice for the number of fuzzy subsets would be a value of 1000. However, for the KNN method, it is important noting that when more datapoints are available, one can choose to use a larger value of k-nearest neighbors. As the results of Table 38 show,

the performance of the KNN method for detecting an increase or decrease in uncertainty with the increase in the number of replications varies considerably based on the number of k-nearest neighbors used and also on the number of replications used, behavior that is not observed from the other methods. In this case, a good choice for k-nearest neighbors appears to be dependent on the number of replications being considered. Nevertheless, it appears that a mid-range value of k-nearest neighbors (something around 50% of the data available) is a good rule of thumb of be used.

Table 37. Results from KNN method using different number of neighbors (k) for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	k neighbors	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	1	32.35%	32.35%	44.16%	54.57%
	2	46.55%	46.55%	44.30%	54.71%
	3	63.15%	63.15%	44.30%	54.71%
	4	73.00%	73.00%	44.30%	54.71%
	5	81.58%	81.58%	44.30%	54.71%
	6	80.45%	80.45%	44.30%	54.71%
	7	80.17%	80.17%	44.30%	54.71%
	8	78.76%	78.76%	44.30%	54.71%
	9	79.47%	79.47%	44.30%	54.71%
$H(X_2)$	1	32.63%	32.63%	45.43%	52.18%
	2	47.12%	47.12%	45.43%	52.18%
	3	64.84%	64.84%	45.43%	52.18%
	4	73.42%	73.42%	45.43%	52.18%
	5	79.75%	79.75%	45.43%	52.18%
	6	78.34%	78.34%	45.43%	52.18%
	7	77.64%	77.64%	45.43%	52.18%
	8	78.90%	78.90%	45.43%	52.18%
	9	79.61%	79.61%	45.43%	52.18%
$H(X_3)$	1	68.33%	68.33%	69.44%	74.44%

	2	74.44%	74.44%	69.44%	74.44%
	3	81.67%	81.67%	69.44%	74.44%
	4	88.89%	88.89%	69.44%	74.44%
	5	91.67%	91.67%	69.44%	74.44%
	6	89.44%	89.44%	69.44%	74.44%
	7	92.22%	92.22%	69.44%	74.44%
	8	92.78%	92.78%	69.44%	74.44%
	9	88.33%	88.33%	69.44%	74.44%
$H(Y_1)$	1	33.19%	33.19%	48.95%	54.43%
	2	48.66%	48.66%	48.95%	54.43%
	3	65.82%	65.82%	48.95%	54.43%
	4	77.22%	77.22%	48.95%	54.43%
	5	81.29%	81.29%	48.95%	54.43%
	6	78.76%	78.76%	49.09%	54.57%
	7	80.17%	80.17%	49.09%	54.57%
	8	80.45%	80.45%	49.09%	54.57%
	9	82.14%	82.14%	49.09%	54.57%
$H(Y_2)$	1	32.49%	32.49%	48.10%	53.31%
	2	47.26%	47.26%	48.10%	53.31%
	3	65.26%	65.26%	48.24%	53.45%
	4	73.56%	73.56%	48.24%	53.45%
	5	78.34%	78.34%	48.24%	53.45%
	6	77.78%	77.78%	48.24%	53.45%
	7	78.06%	78.06%	48.24%	53.45%
	8	78.76%	78.76%	48.24%	53.45%
	9	81.15%	81.15%	48.24%	53.45%

Table 38. $H(X_1)$ results from KNN method using different number of neighbors (k) per different number of replications for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Number of replications	k neighbors	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	20	1	97.47%	97.47%	10.13%	8.86%
		3	100.00%	100.00%	11.39%	10.13%
		5	100.00%	100.00%	11.39%	10.13%
		7	100.00%	100.00%	11.39%	10.13%
		9	100.00%	100.00%	11.39%	10.13%

	11	93.67%	93.67%	11.39%	10.13%
	13	100.00%	100.00%	11.39%	10.13%
	15	100.00%	100.00%	11.39%	10.13%
	17	100.00%	100.00%	11.39%	10.13%
	19	100.00%	100.00%	11.39%	10.13%
50	1	100.00%	100.00%	93.67%	93.67%
	3	73.42%	73.42%	93.67%	93.67%
	5	100.00%	100.00%	93.67%	93.67%
	7	100.00%	100.00%	93.67%	93.67%
	9	100.00%	100.00%	93.67%	93.67%
	13	100.00%	100.00%	93.67%	93.67%
	19	100.00%	100.00%	93.67%	93.67%
	25	100.00%	100.00%	93.67%	93.67%
	35	100.00%	100.00%	93.67%	93.67%
	49	100.00%	100.00%	93.67%	93.67%
100	1	88.61%	88.61%	5.06%	2.53%
	3	93.67%	93.67%	5.06%	2.53%
	5	94.94%	94.94%	5.06%	2.53%
	7	94.94%	94.94%	5.06%	2.53%
	9	100.00%	100.00%	5.06%	2.53%
	13	94.94%	94.94%	5.06%	2.53%
	19	93.67%	93.67%	5.06%	2.53%
	25	94.94%	94.94%	5.06%	2.53%
	49	100.00%	100.00%	5.06%	2.53%
	99	93.67%	93.67%	5.06%	2.53%
200	1	5.06%	5.06%	8.86%	16.46%
	3	89.87%	89.87%	8.86%	16.46%
	7	91.14%	91.14%	8.86%	16.46%
	9	88.61%	88.61%	8.86%	16.46%
	19	93.67%	93.67%	8.86%	16.46%
	25	94.94%	94.94%	8.86%	16.46%
	49	84.81%	84.81%	8.86%	16.46%
	99	53.16%	53.16%	8.86%	16.46%
	150	2.53%	2.53%	8.86%	16.46%
	199	89.87%	89.87%	8.86%	16.46%
400	1	0.00%	0.00%	81.01%	92.41%
	3	88.61%	88.61%	81.01%	92.41%
	9	88.61%	88.61%	81.01%	92.41%
	19	84.81%	84.81%	81.01%	92.41%
	25	88.61%	88.61%	81.01%	92.41%

	49	89.87%	89.87%	81.01%	92.41%
	99	94.94%	94.94%	81.01%	92.41%
	150	98.73%	98.73%	81.01%	92.41%
	199	100.00%	100.00%	81.01%	92.41%
	399	87.34%	87.34%	81.01%	92.41%
600	1	0.00%	0.00%	0.00%	81.01%
	3	13.92%	13.92%	0.00%	81.01%
	9	15.19%	15.19%	0.00%	81.01%
	25	10.13%	10.13%	0.00%	81.01%
	49	2.53%	2.53%	0.00%	81.01%
	99	10.13%	10.13%	0.00%	81.01%
	150	67.09%	67.09%	0.00%	81.01%
	199	2.53%	2.53%	0.00%	81.01%
	399	1.27%	1.27%	0.00%	81.01%
	599	17.72%	17.72%	0.00%	81.01%
800	1	0.00%	0.00%	0.00%	0.00%
	3	45.57%	45.57%	0.00%	0.00%
	9	77.22%	77.22%	0.00%	0.00%
	25	54.43%	54.43%	0.00%	0.00%
	49	44.30%	44.30%	0.00%	0.00%
	99	54.43%	54.43%	0.00%	0.00%
	199	54.43%	54.43%	0.00%	0.00%
	399	25.32%	25.32%	0.00%	0.00%
	599	24.05%	24.05%	0.00%	0.00%
	799	67.09%	67.09%	0.00%	0.00%
1000	1	0.00%	0.00%	100.00%	100.00%
	3	63.29%	63.29%	100.00%	100.00%
	9	91.14%	91.14%	100.00%	100.00%
	25	97.47%	97.47%	100.00%	100.00%
	49	97.47%	97.47%	100.00%	100.00%
	99	98.73%	98.73%	100.00%	100.00%
	199	98.73%	98.73%	100.00%	100.00%
	399	100.00%	100.00%	100.00%	100.00%
	599	100.00%	100.00%	100.00%	100.00%
	999	98.73%	98.73%	100.00%	100.00%
1500	1	0.00%	0.00%	98.73%	96.20%
	3	0.00%	0.00%	98.73%	96.20%
	9	43.04%	43.04%	98.73%	96.20%
	25	93.67%	93.67%	98.73%	96.20%
	49	89.87%	89.87%	98.73%	96.20%

	99	94.94%	94.94%	98.73%	96.20%
	199	94.94%	94.94%	98.73%	96.20%
	599	98.73%	98.73%	98.73%	96.20%
	999	13.92%	13.92%	98.73%	96.20%
	1499	96.20%	96.20%	98.73%	96.20%

Table 39. Results from KNN method using different number of neighbors (k) for detecting the experiment that leads to the maximum uncertainty.

Entropy	k neighbors	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	1	0.00%	0.00%	8.86%	8.86%
	2	0.00%	0.00%	10.13%	10.13%
	3	0.00%	0.00%	10.13%	10.13%
	4	7.59%	7.59%	10.13%	10.13%
	5	34.18%	34.18%	10.13%	10.13%
	6	45.57%	45.57%	10.13%	10.13%
	7	59.49%	59.49%	10.13%	10.13%
	8	26.58%	26.58%	10.13%	10.13%
	9	36.71%	36.71%	10.13%	10.13%
$H(X_2)$	1	0.00%	0.00%	2.53%	0.00%
	2	0.00%	0.00%	2.53%	0.00%
	3	0.00%	0.00%	2.53%	0.00%
	4	15.19%	15.19%	2.53%	0.00%
	5	35.44%	35.44%	2.53%	0.00%
	6	31.65%	31.65%	2.53%	0.00%
	7	36.71%	36.71%	2.53%	0.00%
	8	40.51%	40.51%	2.53%	0.00%
	9	41.77%	41.77%	2.53%	0.00%
$H(X_3)$	1	50.00%	50.00%	50.00%	50.00%
	2	50.00%	50.00%	50.00%	50.00%
	3	50.00%	50.00%	50.00%	50.00%
	4	70.00%	70.00%	50.00%	50.00%
	5	95.00%	95.00%	50.00%	50.00%
	6	85.00%	85.00%	50.00%	50.00%
	7	95.00%	95.00%	50.00%	50.00%
	8	95.00%	95.00%	50.00%	50.00%
	9	95.00%	95.00%	50.00%	50.00%

	1	0.00%	0.00%	13.92%	11.39%
	2	0.00%	0.00%	13.92%	11.39%
	3	1.27%	1.27%	13.92%	11.39%
	4	29.11%	29.11%	13.92%	11.39%
$H(Y_1)$	5	45.57%	45.57%	13.92%	11.39%
	6	43.04%	43.04%	13.92%	11.39%
	7	51.90%	51.90%	13.92%	11.39%
	8	49.37%	49.37%	13.92%	11.39%
	9	55.70%	55.70%	13.92%	11.39%
	1	0.00%	0.00%	12.66%	6.33%
	2	0.00%	0.00%	12.66%	6.33%
	3	0.00%	0.00%	12.66%	6.33%
	4	17.72%	17.72%	12.66%	6.33%
$H(Y_2)$	5	39.24%	39.24%	12.66%	6.33%
	6	37.97%	37.97%	12.66%	6.33%
	7	44.30%	44.30%	12.66%	6.33%
	8	46.84%	46.84%	12.66%	6.33%
	9	51.90%	51.90%	12.66%	6.33%

Table 40. Results from fuzzy-histogram based method with different number of fuzzy subsets and different fuzzy membership functions for detecting an increase or decrease in uncertainty with the increase in the number of replications.

Entropy	Fuzzy membership function	Number of fuzzy subsets	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
		2	12.38%	12.38%	44.30%	54.71%
		5	17.58%	17.58%	44.30%	54.71%
		10	31.79%	31.79%	44.30%	54.71%
		25	46.55%	46.55%	44.30%	54.71%
	cosine	50	57.10%	57.10%	44.30%	54.71%
		100	71.31%	71.31%	44.30%	54.71%
$H(X_1)$		200	87.20%	87.20%	44.30%	54.71%
		500	97.61%	97.61%	44.30%	54.71%
		1000	99.30%	99.30%	44.30%	54.71%
		2	12.24%	12.24%	44.30%	54.71%
		5	21.80%	21.80%	44.30%	54.71%
	crisp	10	31.79%	31.79%	44.30%	54.71%
		25	47.96%	47.96%	44.30%	54.71%

		50	58.23%	58.23%	44.30%	54.71%
		100	75.11%	75.11%	44.30%	54.71%
		200	93.81%	93.81%	44.30%	54.71%
		500	98.59%	98.59%	44.30%	54.71%
		1000	100.00%	100.00%	44.30%	54.71%
	triangular	2	17.86%	17.86%	44.30%	54.71%
		5	16.32%	16.32%	44.30%	54.71%
		10	31.36%	31.36%	44.30%	54.71%
		25	45.71%	45.71%	44.30%	54.71%
		50	54.71%	54.71%	44.30%	54.71%
		100	70.32%	70.32%	44.30%	54.71%
		200	81.01%	81.01%	44.30%	54.71%
		500	97.61%	97.61%	44.30%	54.71%
		1000	99.30%	99.30%	44.30%	54.71%
$H(X_2)$	cosine	2	16.46%	16.46%	45.43%	52.18%
		5	18.00%	18.00%	45.43%	52.18%
		10	28.13%	28.13%	45.43%	52.18%
		25	46.84%	46.84%	45.43%	52.18%
		50	56.12%	56.12%	45.43%	52.18%
		100	72.71%	72.71%	45.43%	52.18%
		200	90.01%	90.01%	45.43%	52.18%
		500	98.03%	98.03%	45.43%	52.18%
		1000	100.00%	100.00%	45.71%	52.46%
	crisp	2	14.21%	14.21%	45.43%	52.18%
		5	24.75%	24.75%	45.43%	52.18%
		10	28.97%	28.97%	45.43%	52.18%
		25	50.91%	50.91%	45.43%	52.18%
		50	65.54%	65.54%	45.43%	52.18%
		100	78.06%	78.06%	45.43%	52.18%
		200	95.78%	95.78%	45.43%	52.18%
		500	99.02%	99.02%	45.43%	52.18%
		1000	100.00%	100.00%	45.43%	52.18%
	triangular	2	18.14%	18.14%	45.43%	52.18%
		5	13.50%	13.50%	45.43%	52.18%
		10	28.27%	28.27%	45.43%	52.18%
		25	45.29%	45.29%	45.43%	52.18%
		50	54.29%	54.29%	45.43%	52.18%
		100	71.03%	71.03%	45.43%	52.18%
		200	86.08%	86.08%	45.43%	52.18%
		500	98.03%	98.03%	45.43%	52.18%

		1000	99.86%	99.86%	45.43%	52.18%
$H(X_3)$	cosine	2	60.00%	60.00%	69.44%	74.44%
		5	59.44%	59.44%	69.44%	74.44%
		10	66.11%	66.11%	69.44%	74.44%
		25	75.00%	75.00%	69.44%	74.44%
		50	79.44%	79.44%	69.44%	74.44%
		100	87.78%	87.78%	69.44%	74.44%
		200	92.22%	92.22%	69.44%	74.44%
		500	98.33%	98.33%	69.44%	74.44%
		1000	100.00%	100.00%	69.44%	74.44%
	crisp	2	61.11%	61.11%	69.44%	74.44%
		5	63.33%	63.33%	69.44%	74.44%
		10	67.22%	67.22%	69.44%	74.44%
		25	75.00%	75.00%	69.44%	74.44%
		50	83.89%	83.89%	69.44%	74.44%
		100	90.56%	90.56%	69.44%	74.44%
		200	96.11%	96.11%	69.44%	74.44%
		500	99.44%	99.44%	69.44%	74.44%
		1000	100.00%	100.00%	69.44%	74.44%
	triangular	2	62.78%	62.78%	69.44%	74.44%
		5	59.44%	59.44%	69.44%	74.44%
		10	65.56%	65.56%	69.44%	74.44%
		25	74.44%	74.44%	69.44%	74.44%
		50	79.44%	79.44%	69.44%	74.44%
		100	86.67%	86.67%	69.44%	74.44%
		200	92.22%	92.22%	69.44%	74.44%
		500	98.33%	98.33%	69.44%	74.44%
		1000	100.00%	100.00%	69.44%	74.44%
$H(Y_1)$	cosine	2	16.32%	16.32%	48.95%	54.43%
		5	20.82%	20.82%	48.95%	54.43%
		10	25.32%	25.32%	48.95%	54.43%
		25	43.60%	43.60%	48.95%	54.43%
		50	57.10%	57.10%	48.95%	54.43%
		100	69.76%	69.76%	48.95%	54.43%
		200	84.67%	84.67%	48.95%	54.43%
		500	96.77%	96.77%	48.95%	54.43%
		1000	98.31%	98.31%	48.95%	54.43%
	crisp	2	17.86%	17.86%	48.95%	54.43%
		5	24.75%	24.75%	48.95%	54.43%
		10	28.27%	28.27%	48.95%	54.43%

		25	46.55%	46.55%	48.95%	54.43%
		50	64.28%	64.28%	48.95%	54.43%
		100	73.70%	73.70%	48.95%	54.43%
		200	91.98%	91.98%	48.95%	54.43%
		500	97.47%	97.47%	48.95%	54.43%
		1000	98.59%	98.59%	48.95%	54.43%
	triangular	2	15.61%	15.61%	48.95%	54.43%
		5	14.77%	14.77%	48.95%	54.43%
		10	25.32%	25.32%	48.95%	54.43%
		25	42.33%	42.33%	48.95%	54.43%
		50	54.29%	54.29%	48.95%	54.43%
		100	67.93%	67.93%	48.95%	54.43%
		200	81.43%	81.43%	48.95%	54.43%
		500	95.92%	95.92%	48.95%	54.43%
		1000	98.31%	98.31%	48.95%	54.43%
$H(Y_2)$	cosine	2	15.61%	15.61%	48.10%	53.31%
		5	18.71%	18.71%	48.10%	53.31%
		10	24.33%	24.33%	48.10%	53.31%
		25	42.48%	42.48%	48.10%	53.31%
		50	54.99%	54.99%	48.10%	53.31%
		100	70.18%	70.18%	48.10%	53.31%
		200	86.78%	86.78%	48.10%	53.31%
		500	97.75%	97.75%	48.10%	53.31%
		1000	99.30%	99.30%	48.24%	53.45%
	crisp	2	15.89%	15.89%	48.10%	53.31%
		5	24.05%	24.05%	48.10%	53.31%
		10	27.99%	27.99%	48.10%	53.31%
		25	46.69%	46.69%	48.10%	53.31%
		50	61.88%	61.88%	48.10%	53.31%
		100	74.12%	74.12%	48.10%	53.31%
		200	93.95%	93.95%	48.10%	53.31%
		500	98.59%	98.59%	48.10%	53.31%
		1000	99.44%	99.44%	48.10%	53.31%
	triangular	2	16.60%	16.60%	48.10%	53.31%
		5	14.77%	14.77%	48.10%	53.31%
		10	24.19%	24.19%	48.10%	53.31%
		25	40.23%	40.23%	48.10%	53.31%
		50	52.04%	52.04%	48.10%	53.31%
		100	68.35%	68.35%	48.10%	53.31%
		200	82.98%	82.98%	48.10%	53.31%

		500	97.61%	97.61%	48.10%	53.31%
		1000	99.02%	99.02%	48.10%	53.31%

Table 41. Results from fuzzy-histogram based method with different number of fuzzy subsets and different fuzzy membership functions for detecting the experiment that leads to the maximum uncertainty.

Entropy	Fuzzy membership function	Number of fuzzy subsets	Mean of comparison entropy vs. SAE	Mean of comparison entropy vs. SSE	Mean of comparison average entropy vs. MAE	Mean of comparison average entropy vs. MSE
$H(X_1)$	cosine	2	0.00%	0.00%	10.13%	10.13%
		5	0.00%	0.00%	10.13%	10.13%
		10	0.00%	0.00%	10.13%	10.13%
		25	0.00%	0.00%	10.13%	10.13%
		50	0.00%	0.00%	10.13%	10.13%
		100	5.06%	5.06%	10.13%	10.13%
		200	34.18%	34.18%	10.13%	10.13%
		500	93.67%	93.67%	10.13%	10.13%
		1000	100.00%	100.00%	10.13%	10.13%
	crisp	2	0.00%	0.00%	10.13%	10.13%
		5	0.00%	0.00%	10.13%	10.13%
		10	0.00%	0.00%	10.13%	10.13%
		25	0.00%	0.00%	10.13%	10.13%
		50	0.00%	0.00%	10.13%	10.13%
		100	5.06%	5.06%	10.13%	10.13%
		200	73.42%	73.42%	10.13%	10.13%
		500	100.00%	100.00%	10.13%	10.13%
		1000	100.00%	100.00%	10.13%	10.13%
	triangular	2	0.00%	0.00%	10.13%	10.13%
		5	0.00%	0.00%	10.13%	10.13%
		10	0.00%	0.00%	10.13%	10.13%
		25	0.00%	0.00%	10.13%	10.13%
		50	0.00%	0.00%	10.13%	10.13%
		100	2.53%	2.53%	10.13%	10.13%
		200	24.05%	24.05%	10.13%	10.13%
		500	93.67%	93.67%	10.13%	10.13%
		1000	100.00%	100.00%	10.13%	10.13%
$H(X_2)$	cosine	2	0.00%	0.00%	2.53%	0.00%
		5	0.00%	0.00%	2.53%	0.00%

		10	0.00%	0.00%	2.53%	0.00%
		25	0.00%	0.00%	2.53%	0.00%
		50	1.27%	1.27%	2.53%	0.00%
		100	1.27%	1.27%	2.53%	0.00%
		200	55.70%	55.70%	2.53%	0.00%
		500	88.61%	88.61%	2.53%	0.00%
		1000	100.00%	100.00%	2.53%	0.00%
	crisp	2	0.00%	0.00%	2.53%	0.00%
		5	0.00%	0.00%	2.53%	0.00%
		10	0.00%	0.00%	2.53%	0.00%
		25	0.00%	0.00%	2.53%	0.00%
		50	1.27%	1.27%	2.53%	0.00%
		100	12.66%	12.66%	2.53%	0.00%
		200	78.48%	78.48%	2.53%	0.00%
		500	97.47%	97.47%	2.53%	0.00%
		1000	100.00%	100.00%	2.53%	0.00%
	triangular	2	0.00%	0.00%	2.53%	0.00%
		5	0.00%	0.00%	2.53%	0.00%
		10	0.00%	0.00%	2.53%	0.00%
		25	0.00%	0.00%	2.53%	0.00%
		50	1.27%	1.27%	2.53%	0.00%
		100	1.27%	1.27%	2.53%	0.00%
		200	39.24%	39.24%	2.53%	0.00%
		500	88.61%	88.61%	2.53%	0.00%
		1000	98.73%	98.73%	2.53%	0.00%
$H(X_3)$	cosine	2	50.00%	50.00%	50.00%	50.00%
		5	50.00%	50.00%	50.00%	50.00%
		10	50.00%	50.00%	50.00%	50.00%
		25	50.00%	50.00%	50.00%	50.00%
		50	50.00%	50.00%	50.00%	50.00%
		100	55.00%	55.00%	50.00%	50.00%
		200	60.00%	60.00%	50.00%	50.00%
		500	85.00%	85.00%	50.00%	50.00%
		1000	100.00%	100.00%	50.00%	50.00%
	crisp	2	50.00%	50.00%	50.00%	50.00%
		5	50.00%	50.00%	50.00%	50.00%
		10	50.00%	50.00%	50.00%	50.00%
		25	50.00%	50.00%	50.00%	50.00%
		50	50.00%	50.00%	50.00%	50.00%
		100	60.00%	60.00%	50.00%	50.00%

		200	70.00%	70.00%	50.00%	50.00%
		500	95.00%	95.00%	50.00%	50.00%
		1000	100.00%	100.00%	50.00%	50.00%
		2	50.00%	50.00%	50.00%	50.00%
		5	50.00%	50.00%	50.00%	50.00%
		10	50.00%	50.00%	50.00%	50.00%
		25	50.00%	50.00%	50.00%	50.00%
	triangular	50	50.00%	50.00%	50.00%	50.00%
		100	55.00%	55.00%	50.00%	50.00%
		200	60.00%	60.00%	50.00%	50.00%
		500	85.00%	85.00%	50.00%	50.00%
		1000	100.00%	100.00%	50.00%	50.00%
		2	0.00%	0.00%	13.92%	11.39%
		5	0.00%	0.00%	13.92%	11.39%
		10	0.00%	0.00%	13.92%	11.39%
		25	0.00%	0.00%	13.92%	11.39%
	cosine	50	0.00%	0.00%	13.92%	11.39%
		100	1.27%	1.27%	13.92%	11.39%
		200	27.85%	27.85%	13.92%	11.39%
		500	89.87%	89.87%	13.92%	11.39%
		1000	96.20%	96.20%	13.92%	11.39%
		2	0.00%	0.00%	13.92%	11.39%
		5	0.00%	0.00%	13.92%	11.39%
		10	0.00%	0.00%	13.92%	11.39%
		25	0.00%	0.00%	13.92%	11.39%
$H(Y_1)$	crisp	50	0.00%	0.00%	13.92%	11.39%
		100	6.33%	6.33%	13.92%	11.39%
		200	60.76%	60.76%	13.92%	11.39%
		500	94.94%	94.94%	13.92%	11.39%
		1000	96.20%	96.20%	13.92%	11.39%
		2	0.00%	0.00%	13.92%	11.39%
		5	0.00%	0.00%	13.92%	11.39%
		10	0.00%	0.00%	13.92%	11.39%
		25	0.00%	0.00%	13.92%	11.39%
	triangular	50	0.00%	0.00%	13.92%	11.39%
		100	1.27%	1.27%	13.92%	11.39%
		200	20.25%	20.25%	13.92%	11.39%
		500	83.54%	83.54%	13.92%	11.39%
		1000	96.20%	96.20%	13.92%	11.39%
$H(Y_2)$	cosine	2	0.00%	0.00%	12.66%	6.33%

	5	0.00%	0.00%	12.66%	6.33%
	10	0.00%	0.00%	12.66%	6.33%
	25	0.00%	0.00%	12.66%	6.33%
	50	0.00%	0.00%	12.66%	6.33%
	100	5.06%	5.06%	12.66%	6.33%
	200	50.63%	50.63%	12.66%	6.33%
	500	89.87%	89.87%	12.66%	6.33%
	1000	97.47%	97.47%	12.66%	6.33%
crisp	2	0.00%	0.00%	12.66%	6.33%
	5	0.00%	0.00%	12.66%	6.33%
	10	0.00%	0.00%	12.66%	6.33%
	25	0.00%	0.00%	12.66%	6.33%
	50	1.27%	1.27%	12.66%	6.33%
	100	7.59%	7.59%	12.66%	6.33%
	200	72.15%	72.15%	12.66%	6.33%
	500	94.94%	94.94%	12.66%	6.33%
	1000	97.47%	97.47%	12.66%	6.33%
triangular	2	0.00%	0.00%	12.66%	6.33%
	5	0.00%	0.00%	12.66%	6.33%
	10	0.00%	0.00%	12.66%	6.33%
	25	0.00%	0.00%	12.66%	6.33%
	50	0.00%	0.00%	12.66%	6.33%
	100	3.80%	3.80%	12.66%	6.33%
	200	31.65%	31.65%	12.66%	6.33%
	500	89.87%	89.87%	12.66%	6.33%
	1000	96.20%	96.20%	12.66%	6.33%

A final analysis performed in terms of entropy measures involved using χ^2 test to investigate whether the performance of each one the methods used to calculate the entropy measures was statistically significantly different than the other method. The performance referred here is in comparison to the measures of error SAE, SSE, MAE, and MSE. The χ^2 test was performed using JMP® at an α-level of 0.05, where the null hypothesis H_0 means that there is no difference between the performance of the methods.

The results of the χ^2 test are shown in Table 42 to Table 47. As one would expect from the results of each method presented in Table 33 to

Table 41, the methods are statistically significantly different from each other when they are compared against the SSE or SAE measure of error. However, when the methods are compared against either the MAE or MSE measure of error, the χ^2 test did not have enough evidence that one of the methods was statistically significant different than the other. A possible explanation for the lack of evidence to reject H_0 is that the results of the methods when compared to the MAE and MSE measures of error did not vary considerably.

The results of the methods compared to the SAE measure of error are not presented here because they are identical to the results of the methods compared to the SSE measure of error, as indicated on Table 42 and Table 45 below.

Table 42. χ^2 test comparing the performance of the methods with respect to their capability of detecting an increase or decrease in uncertainty with the increase in the number of replications in agreement with the SSE measure of error (or SAE measure of error).

Entropy	Proportion difference			P-value		
	Fuzzy vs KNN	**Kernel vs KNN**	**Kernel vs Fuzzy**	**Fuzzy vs KNN**	**Kernel vs KNN**	**Kernel vs Fuzzy**
$H(X_1)$	0.1007	0.0369	0.1376	**<0.0001**	**<0.0001**	**<0.0001**
$H(X_2)$	0.0884	0.0261	0.1144	**<0.0001**	**0.0004**	**<0.0001**
$H(X_3)$	0.0479	0.2958	0.2479	**0.0004**	**<0.0001**	**<0.0001**
$H(Y_1)$	0.1226	0.0289	0.1516	**<0.0001**	**<0.0001**	**<0.0001**
$H(Y_2)$	0.1072	0.0354	0.1426	**<0.0001**	**<0.0001**	**<0.0001**

Table 43. χ^2 test comparing the performance of the methods with respect to their capability of detecting an increase or decrease in uncertainty with the increase in the number of replications in agreement with the MAE measure of error.

Entropy	Proportion difference			P-value		
	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy
$H(X_1)$	0.0002	0.0002	0.0000	0.9997	0.9997	1.0000
$H(X_2)$	0.0001	0.0000	0.0001	0.9999	1.0000	0.9998
$H(X_3)$	0.0000	0.5000	0.5000	1.0000	**<0.0001**	**<0.0001**
$H(Y_1)$	0.0006	0.0011	0.0017	0.9959	0.9885	0.9433
$H(Y_2)$	0.0010	0.0001	0.0009	0.9886	0.9998	0.9825

Table 44. χ^2 test comparing the performance of the methods with respect to their capability of detecting an increase or decrease in uncertainty with the increase in the number of replications in agreement with the MSE measure of error.

Entropy	Proportion difference			P-value		
	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy
$H(X_1)$	0.0002	0.0002	0.0000	0.9997	0.9997	1.0000
$H(X_2)$	0.0001	0.0000	0.0001	0.9999	1.0000	0.9998
$H(X_3)$	0.0000	0.5000	0.5000	1.0000	**<0.0001**	**<0.0001**
$H(Y_1)$	0.0006	0.0011	0.0017	0.9958	0.9884	0.9428
$H(Y_2)$	0.0010	0.0001	0.0009	0.9885	0.9998	0.9824

Table 45. χ^2 test comparing the performance of the methods with respect to their capability of detecting the experiment that leads to the maximum uncertainty in agreement with the SSE measure of error (or SAE measure of error).

Entropy	Proportion difference			P-value		
	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy
$H(X_1)$	0.0375	0.4433	0.4057	0.1336	**<0.0001**	**<0.0001**
$H(X_2)$	0.0600	0.4838	0.4238	**0.0056**	**<0.0001**	**<0.0001**
$H(X_3)$	0.1407	0.2976	0.1569	**0.0021**	**<0.0001**	**<0.0001**
$H(Y_1)$	0.0567	0.4022	0.4590	**0.0094**	**<0.0001**	**<0.0001**
$H(Y_2)$	0.0089	0.4659	0.4570	0.8886	**<0.0001**	**<0.0001**

Table 46. χ^2 test comparing the performance of the methods with respect to their capability of detecting the experiment that leads to the maximum uncertainty in agreement with the MAE measure of error.

Entropy	Proportion difference			P-value		
	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy
$H(X_1)$	0.0014	0.0014	0.0000	0.9936	0.9937	1.0000
$H(X_2)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000
$H(X_3)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000
$H(Y_1)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000
$H(Y_2)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000

Table 47. χ^2 test comparing the performance of the methods with respect to their capability of detecting the experiment that leads to the maximum uncertainty in agreement with the MSE measure of error.

Entropy	Proportion difference			P-value		
	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy	Fuzzy vs KNN	Kernel vs KNN	Kernel vs Fuzzy
$H(X_1)$	0.0014	0.0014	0.0000	0.9936	0.9937	1.0000
$H(X_2)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000
$H(X_3)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000
$H(Y_1)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000
$H(Y_2)$	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000

Mutual information quantifies the average reduction in uncertainty in the value of Y provided by the value of X and vice-versa and it is also a measure of dependence between X and Y. In order to investigate the potential of the MI as a method of uncertainty quantification in simulation model, the MI results were compared to the results of three other measures of dependence between variables, namely: distance correlation, Pearson correlation, and ${R^2}_{adj}$. To perform these comparisons the procedure described in section 2.4.4 was followed and the results are shown in Table 77 to Table 82 of the Appendix. Table 77 and Table 78 show the results for the MI calculated using the kernel method.

Table 79 and Table 80 show the results for the MI calculated using the KNN method. Table 81 and Table 82 show the results for the MI calculated using the fuzzy-histogram based method.

Table 77 to Table 82 of the Appendix show the results of the comparisons at a granular level. For an easier and more objective comparison, Table 48 shows the results summarized by calculation method, measure of dependence, and impact on the output. Table 49 also summarizes the results by calculation method, measure of dependence, and impact on the output, and it includes the function type (kernel function or fuzzy membership function) as well. As the results in Table 48 show, overall the calculation methods showed better performance when compared with the distance correlation measure than when compared with Pearson correlation or ${R^2}_{adj}$. This is expected as distance correlation is the measure of dependence that is able to capture both linear and non-linear relation between variables similarly to MI. Among the methods, the KNN method led to the best performance and the fuzzy-histogram method led to the worst performance. KNN was not the best method when considering the greatest impact on the TIS. In this case, the kernel method showed the best performance.

Table 49 does not include results for the KNN method because this method does not use any function on its calculation. As can be seen in Table 49, the kernel function or fuzzy membership function used to calculate the MI measures does not appear to have a significant impact on the performance of the MI because the results do not vary considerably by function type. However, this need to be statistically verified and a logistic regression model was used for that, which is discussed next.

Table 48. Results from the comparison of the MI vs the measures of dependence summarized by calculation method, measure of dependence, and impact on the output.

Calculation Method	Measure of dependence	Greatest impact on		Least impact on	
		NIS	TIS	NIS	TIS
Kernel	Distance correlation	43.35%	43.99%	43.02%	43.45%
	Pearson correlation	36.07%	43.01%	36.38%	43.13%
	R^2_{adj}	36.60%	43.28%	36.71%	43.50%
KNN	Distance correlation	45.90%	40.26%	50.82%	47.33%
	Pearson correlation	39.80%	38.84%	43.56%	45.39%
	R^2_{adj}	39.91%	38.76%	43.08%	44.59%
Fuzzy-histogram based	Distance correlation	40.16%	25.38%	43.60%	31.91%
	Pearson correlation	34.26%	19.26%	37.03%	25.30%
	R^2_{adj}	34.29%	19.23%	36.52%	24.50%

Table 49. Results from the comparison of the MI vs the measures of dependence summarized by calculation method, measure of dependence, function, and impact on the output.

Calculation Method	Measure of dependence	Function (kernel or membership function)	Greatest impact on		Least impact on	
			NIS	TIS	NIS	TIS
Kernel	Distance correlation	Epanechnikov	43.06%	43.68%	42.74%	43.14%
		Normal	43.64%	44.30%	43.30%	43.77%
	Pearson correlation	Epanechnikov	35.81%	42.72%	36.15%	42.85%
		Normal	36.32%	43.31%	36.61%	43.42%
	R^2_{adj}	Epanechnikov	36.34%	42.98%	36.49%	43.24%
		Normal	36.85%	43.57%	36.92%	43.76%
Fuzzy-histogram based	Distance correlation	Cosine	40.02%	25.57%	43.41%	32.06%
		Crisp	40.79%	25.32%	44.21%	31.88%
		Triangular	39.69%	25.24%	43.19%	31.80%
	Pearson correlation	Cosine	34.13%	19.22%	36.85%	25.24%
		Crisp	35.34%	20.01%	38.10%	26.04%
		Triangular	33.31%	18.53%	36.15%	24.62%
	R^2_{adj}	Cosine	34.12%	19.20%	36.34%	24.44%
		Crisp	35.44%	19.97%	37.58%	25.24%
		Triangular	33.31%	18.52%	35.62%	23.82%

Logistic regression with two-factor interaction effect was performed in JMP® to investigate which factors affected the performance of the MI when compared against the different measures of dependence. Next, χ^2 test was performed in JMP® to investigate whether the performance of MI was statistically significantly different based on the method used to calculate the MI.

Logistic regression was performed on the results of MI compared with each measure of dependence (distance correlation, Pearson correlation, and ${R^2}_{adj}$), from each calculation method (kernel, KNN, and fuzzy-histogram based method), and from each combination of greatest or least impact on the simulation output (input with the greatest impact on the NIS, input with the least impact on the NIS, input with the greatest impact on the TIS, and input with the least impact on the TIS). For the logistic regression, the following factors were considered as possible dependent variables that could affect the performance of the MI (the independent variable): (i) the value of bandwidth, the number of k-nearest neighbors, or the number of fuzzy subsets for the kernel, KNN, or fuzzy-histogram based method, respectively, (ii) the number of replications (10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500), and (iii) the function type. The function type was only considered for the kernel and fuzzy-histogram based methods. In the kernel method, the function type was the kernel function, which was either normal or Epanechnikov. For the fuzzy-histogram based method, the function type was the fuzzy membership function, which was either cosine, crisp, or triangular. It is also important to mention that when Silverman bandwidth was used, the bandwidth value was not a factor to be considered as a possible dependent variable because there was only one sample in this case.

Based on the whole model test, which compares the whole-model fit to the model that omits all the logistic regression parameters except the intercepts, the model considered here was statistically a better fit than the intercepts (p-value less than 0.0001 at α-level of 0.05) for each one of the results of Table 50. The null hypothesis is that the model fits no better than the model that includes only the intercepts. Table 50 shows the misclassification rate of the logistic model by method of calculation, measure of dependence, and impact on the simulation output. Considering only the aforementioned factors as variables that impact the performance of the MI the following average misclassification rate was obtained: 41.04% for the kernel method, 43.05% for the KNN method, and 30.95% for the fuzzy-histogram based method.

Table 50. Misclassification rate of the logistic regression model by method of calculation, measure of dependence, and impact on simulation output.

Method of calculation	Measure of dependence	Impact on simulation output	Misclassification rate
Kernel	Distance correlation	Greatest impact on NIS	0.4335
Kernel	Distance correlation	Least impact on NIS	0.4302
Kernel	Distance correlation	Greatest impact on TIS	0.4399
Kernel	Distance correlation	Least impact on TIS	0.4345
Kernel	Pearson correlation	Greatest impact on NIS	0.3607
Kernel	Pearson correlation	Least impact on NIS	0.3638
Kernel	Pearson correlation	Greatest impact on TIS	0.4301
Kernel	Pearson correlation	Least impact on TIS	0.4313
Kernel	R^2_{adj} R^2_{adj}	Greatest impact on NIS	0.3660

Kernel	R^2_{adj}	Least impact on NIS	0.3671
Kernel	R^2_{adj}	Greatest impact on TIS	0.4328
Kernel	R^2_{adj}	Least impact on TIS	0.4350
KNN	Distance correlation	Greatest impact on NIS	0.4590
KNN	Distance correlation	Least impact on NIS	0.4918
KNN	Distance correlation	Greatest impact on TIS	0.4026
KNN	Distance correlation	Least impact on TIS	0.4733
KNN	Pearson correlation	Greatest impact on NIS	0.3980
KNN	Pearson correlation	Least impact on NIS	0.4356
KNN	Pearson correlation	Greatest impact on TIS	0.3884
KNN	Pearson correlation	Least impact on TIS	0.4539
KNN	R^2_{adj} R^2_{adj}	Greatest impact on NIS	0.3991
KNN	R^2_{adj}	Least impact on NIS	0.4308
KNN	R^2_{adj}	Greatest impact on TIS	0.3876
KNN	R^2_{adj}	Least impact on TIS	0.4459
Fuzzy-histogram	Distance correlation	Greatest impact on NIS	0.4016
Fuzzy-histogram	Distance correlation	Least impact on NIS	0.4360
Fuzzy-histogram	Distance correlation	Greatest impact on TIS	0.2538
Fuzzy-histogram	Distance correlation	Least impact on TIS	0.3191
Fuzzy-histogram	Pearson correlation	Greatest impact on NIS	0.3426
Fuzzy-histogram	Pearson correlation	Least impact on NIS	0.3703
Fuzzy-histogram	Pearson correlation	Greatest impact on TIS	0.1926
Fuzzy-histogram	Pearson correlation	Least impact on TIS	0.2530
Fuzzy-histogram	R^2_{adj} R^2_{adj}	Greatest impact on NIS	0.3429
Fuzzy-histogram	R^2_{adj}	Least impact on NIS	0.3652

Fuzzy-histogram	$R^2{}_{adj}$	Greatest impact on TIS	0.1923
Fuzzy-histogram	$R^2{}_{adj}$	Least impact on TIS	0.2450

Table 51 shows the p-value and order of importance for the factors and effects in the logistic regression model for an α-level of 0.05. As shown in Table 51, in general number of replications appears to be the most important factor in every method. Overall, number of replications and bandwidth value (or number of k-nearest neighbors or number of fuzzy subsets), and the interaction among these factors were the most important factors in the logistic model. In the fuzzy-histogram based model, the fuzzy membership function also appeared to be important in a few cases, especially when the MI was compared against the Pearson correlation or the $R^2{}_{adj}$. These results come to an agreement with the observations initially made from Table 49.

Table 51. P-value and order of importance of factors on logistic regression model.

Method of calculation	Bandwidth type	Measure of dependence	Factor / P-value (order of importance)	Greatest impact on NIS	Least impact on NIS	Greatest impact on TIS	Least impact on TIS
Kernel	Silverman bandwidth	Distance correlation	Number of replications	**0.0034 (1)**	**0.0012 (1)**	0.2788 (1)	**0.0353 (1)**
			Function type	0.9961 (2)	0.8279 (2)	0.9668 (2)	0.8669 (2)
			Number of replications x function type	0.9965 (3)	0.9987 (3)	0.9990 (3)	0.9988 (3)
		Pearson correlation	Number of replications	**0.0146 (1)**	**0.0001 (1)**	**0.0236 (1)**	**0.0004 (1)**
			Function type	0.9808 (2)	0.8113 (2)	0.9951 (2)	0.9366 (2)
			Number of replications x function type	0.9978 (3)	0.9984 (3)	0.9992 (3)	0.9990 (3)

	R^2_{adj}	Number of replications	**0.0019 (1)**	**0.0000 (1)**	**0.0149 (1)**	**0.0001 (1)**
		Function type	0.9769 (2)	0.8130 (2)	0.9948 (2)	0.8689 (2)
		Number of replications x function type	0.9979 (3)	0.9984 (3)	0.9992 (3)	0.9971 (3)
Different values of bandwidth	Distance correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (3)**	**0.0000 (3)**
		Value of bandwidth	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (1)**	**0.0000 (1)**
		Function type	0.1929 (4)	0.1891 (4)	0.1566 (4)	0.1411 (4)
		Number of replications x value of bandwidth	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (2)**
		Number of replications x function type	0.8422 (6)	0.8498 (6)	0.8824 (6)	0.8776 (6)
		Value of bandwidth x function type	0.8108 (5)	0.8261 (5)	0.7721 (5)	0.7013 (5)
	Pearson correlation	Number of replications	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (3)**	**0.0000 (3)**
		Value of bandwidth	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (1)**	**0.0000 (1)**
		Function type	0.2691 (4)	0.3142 (4)	0.1816 (4)	0.2156 (4)
		Number of replications x value of bandwidth	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (2)**
		Number of replications x function type	0.9146 (6)	0.8150 (6)	0.9288 (6)	0.9370 (6)
		Value of bandwidth x function type	0.7246 (5)	0.9111 (5)	0.8254 (5)	0.8261 (5)
	R^2_{adj}	Number of replications	**0.0000 (2)**	**0.0000 (2)**	**0.0000 (3)**	**0.0000 (3)**
		Value of bandwidth	**0.0000 (3)**	**0.0000 (3)**	**0.0000 (1)**	**0.0000 (1)**
		Function type	0.2853 (4)	0.3476 (4)	0.1921 (3)	0.2504 (4)
		Number of replications x value of bandwidth	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (2)**
		Number of replications x function type	0.9260 (6)	0.9342 (6)	0.9402 (6)	0.9629 (6)

		Value of bandwidth x function type	0.7443 (5)	0.8188 (5)	0.8378 (5)	0.8414 (5)
KNN	Distance correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of k-nearest neighbors	0.9109 (2)	0.7124 (2)	**0.0000 (2)**	**0.0001 (2)**
		Number of replications x number of k-nearest neighbors	0.9764 (3)	0.9885 (3)	**0.0001 (3)**	**0.0067 (3)**
	Pearson correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of k-nearest neighbors	0.2501 (2)	0.5266 (2)	**0.0022 (3)**	**0.0020 (3)**
		Number of replications x number of k-nearest neighbors	0.5973 (3)	0.8861 (3)	**0.0001 (2)**	**0.0006 (2)**
	$R^2{}_{adj}$	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of k-nearest neighbors	0.2732 (2)	0.5809 (2)	**0.0021 (3)**	**0.0020 (3)**
		Number of replications x number of k-nearest neighbors	0.6130 (3)	0.8690 (3)	**0.0001 (2)**	**0.0006 (2)**
Fuzzy-histogram based	Distance correlation	Number of replications	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**	**0.0000 (1)**
		Number of fuzzy subsets	**0.0000 (2)**	0.5604 (5)	**0.0000 (2)**	**0.0000 (2)**
		Membership function	0.2955 (5)	0.3621 (3)	0.7181 (4)	0.8804 (4)
		Number of replications x number of fuzzy subsets	**0.0331 (4)**	0.3770 (4)	0.9968 (6)	0.3637 (3)
		Number of replications x membership function	0.9037 (6)	0.6378 (6)	**0.0043 (3)**	1.0000 (6)
		Number of fuzzy subsets x Membership function	**0.0021 (3)**	**0.0079 (2)**	0.9935 (5)	0.9999 (5)

Pearson correlation	Number of replications	**0.0000 (2)**	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (1)**
	Number of fuzzy subsets	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (1)**	**0.0000 (2)**
	Membership function	0.0015 (5)	**0.0094 (4)**	**0.0299 (4)**	0.0838 (3)
	Number of replications x number of fuzzy subsets	**0.0001 (4)**	**0.0271 (5)**	**0.0005 (3)**	0.8736 (4)
	Number of replications x membership function	0.9895 (6)	0.9682 (6)	0.1988 (5)	0.9649 (5)
	Number of fuzzy subsets x Membership function	**0.0000 (3)**	**0.0027 (3)**	0.6140 (6)	0.9933 (6)
$R^2{}_{adj}$	Number of replications	**0.0000 (2)**	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (1)**
	Number of fuzzy subsets	**0.0000 (1)**	**0.0000 (2)**	**0.0000 (1)**	**0.0000 (2)**
	Membership function	**0.0010 (5)**	**0.0087 (4)**	**0.0371 (4)**	0.1221 (3)
	Number of replications x number of fuzzy subsets	**0.0001 (4)**	**0.0341 (5)**	**0.0004 (3)**	0.7573 (4)
	Number of replications x membership function	0.9971 (6)	0.9599 (6)	0.1472 (5)	0.9524 (5)
	Number of fuzzy subsets x Membership function	**0.0001 (3)**	**0.0023 (3)**	0.6127 (6)	0.9926 (6)

Finally, χ^2 test at an α-level of 0.05 was performed to investigate whether the performance of the MI was statistically significantly different based on the method used to calculate the MI, as shown in Table 52. The null hypothesis is that the MI performance based on two calculation methods are not different and the alternative hypothesis is that the MI performance based on two calculation methods are different. As can be seen from

Table 52, with three exceptions the methods are statistically significantly different than each other. The results of Table 52 also come to an agreement with the observations made from Table 48. KNN is statistically significantly better than the fuzzy-histogram method in every case and it is also statistically significantly better than the kernel in almost every case. The exception is when considering the greatest impact on the TIS. In this case, the kernel method is either statistically significantly better than the KNN method or there is no evidence that the methods are different. In general, the kernel method is statistically significantly better than the fuzzy-histogram based method.

Table 52. χ^2 test results whether the performance of the MI is statistically significantly different based on the calculation method.

Measure of dependence	Impact on output	Proportion difference	P-value
Distance correlation	Greatest impact on NIS	-0.0268 (Kernel-KNN)	**<0.0001**
	Least impact on NIS	-0.0797 (Kernel-KNN)	**<0.0001**
	Greatest impact on TIS	0.0344 (Kernel-KNN)	**<0.0001**
	Least impact on TIS	-0.0418 (Kernel-KNN)	**<0.0001**
Pearson correlation	Greatest impact on NIS	-0.0385 (Kernel-KNN)	**<0.0001**
	Least impact on NIS	-0.0731 (Kernel-KNN)	**<0.0001**
	Greatest impact on TIS	0.0391 (Kernel-KNN)	1.0000
	Least impact on TIS	-0.0254 (Kernel-KNN)	**<0.0001**
R^2_{adj}	Greatest impact on NIS	-0.0342 (Kernel-KNN)	**<0.0001**
	Least impact on NIS	-0.0651 (Kernel-KNN)	**<0.0001**
	Greatest impact on TIS	0.0425 (Kernel-KNN)	**<0.0001**
	Least impact on TIS	-0.0137 (Kernel-KNN)	**0.0090**
Distance correlation	Greatest impact on NIS	0.0306 (Kernel-Fuzzy)	**<0.0001**
	Least impact on NIS	-0.0075 (Kernel-Fuzzy)	0.9655
	Greatest impact on TIS	0.1832 (Kernel-Fuzzy)	**<0.0001**
	Least impact on TIS	0.1123 (Kernel-Fuzzy)	**<0.0001**
Pearson correlation	Greatest impact on NIS	0.0170 (Kernel-Fuzzy)	**<0.0001**
	Least impact on NIS	-0.0078 (Kernel-Fuzzy)	**<0.0001**

	Greatest impact on TIS	0.2349 (Kernel-Fuzzy)	**<0.0001**
	Least impact on TIS	0.1755 (Kernel-Fuzzy)	**<0.0001**
R^2_{adj}	Greatest impact on NIS	0.0220 (Kernel-Fuzzy)	**<0.0001**
	Least impact on NIS	0.0006 (Kernel-Fuzzy)	0.4456
	Greatest impact on TIS	0.2378 (Kernel-Fuzzy)	**<0.0001**
	Least impact on TIS	0.1872 (Kernel-Fuzzy)	**<0.0001**
Distance correlation	Greatest impact on NIS	-0.0574 (Fuzzy-KNN)	**<0.0001**
	Least impact on NIS	-0.0721 (Fuzzy-KNN)	**<0.0001**
	Greatest impact on TIS	-0.1488 (Fuzzy-KNN)	**<0.0001**
	Least impact on TIS	-0.1541 (Fuzzy-KNN)	**<0.0001**
Pearson correlation	Greatest impact on NIS	-0.0554 (Fuzzy-KNN)	**<0.0001**
	Least impact on NIS	-0.0653 (Fuzzy-KNN)	**<0.0001**
	Greatest impact on TIS	-0.1958 (Fuzzy-KNN)	**<0.0001**
	Least impact on TIS	-0.2009 (Fuzzy-KNN)	**<0.0001**
R^2_{adj}	Greatest impact on NIS	-0.0562 (Fuzzy-KNN)	**<0.0001**
	Least impact on NIS	-0.0656 (Fuzzy-KNN)	**<0.0001**
	Greatest impact on TIS	-0.1953 (Fuzzy-KNN)	**<0.0001**
	Least impact on TIS	-0.2009 (Fuzzy-KNN)	**<0.0001**

3.4.9. Overall comparison

For easy reference and comparison, the results of section 2, which are directly related to this section, are summarized in Table 53. Similarly, the results of this section are summarized next in Table 54.

Table 53. Summary of results investigating information theory as a method for uncertainty quantification in simulation models using the histogram-based method and data normalization as solution for the challenge encountered while applying entropy for continuous variables.

Method / Characteristics and results of the method	**Histogram-based with probability density function and fixed number of bins**	**Histogram-based with probability density function and optimum number of bins**	**Histogram-based with discrete empirical distribution and fixed bins**	**Histogram-based with discrete empirical distribution, fixed bins, and entropy**

				and MI normalization
Characteristics of the method investigated	Number of bins	Optimum number of bins rules (Sturges', Scott's, and FD's rules)	Without normalization	Normalization method for entropy and for MI (arith, geom, joint, theor)
Number of bins or binwidth	Entropy and MI decrease with decrease in binwidth (increase in number of bins)	NA	Entropy and MI increase with decrease in binwidth (increase in number of bins)	Entropy and MI decrease with decrease in binwidth (increase in number of bins) for entropy and MI_{theor}. For other MI normalizations, the MI increase with decrease in binwidth (increase in number of bins)
Bandwidth value	NA	NA	NA	NA
Number of k-nearest neighbors	NA	NA	NA	NA
Number of fuzzy subsets	NA	NA	NA	NA
Number of replications	Entropy and MI increase with increase in number of replications (for lower bins, the curve tends to stabilize quicker)	Entropy and MI increase with increase in number of replications	Entropy and MI increase with increase in number of replications (for lower bins, the curve tends to stabilize quicker)	Entropy and MI increase with increase in number of replications (for lower bins, the curve tends to stabilize quicker)
Different seeds used	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"
Different parameter values	Different entropy and MI when using different parameter values	Different entropy and MI when using different parameter values	Different entropy and MI when using different parameter values	Different entropy and MI when using different parameter values

Different traffic intensities	Not able to identify any pattern with changes in traffic intensity	Not able to identify any pattern with changes in traffic intensity	Not able to identify any pattern with changes in traffic intensity	Not able to identify any pattern with changes in traffic intensity
CONWIP system	Constant entropy and MI for NIS, which tend to go to 0 as number of bins increase	Entropy and MI of NIS slightly decrease with increase in number of replications	Entropy and MI for NIS are equal to 0 regardless of number of bins	Entropy and MI for NIS are equal to 0 regardless of number of bins. However, for MI_{joint}, MI_{arith}, and MI_{geom}, MI for TIS is equal or close to 1.
Travel time deterministic	Constant entropy and MI for travel time deterministic, which tend to go to 0 as number of bins increase	Entropy and MI of travel time deterministic slightly decrease with increase in number of replications	Entropy and MI for travel time deterministic are equal to 0 regardless of number of bins	Entropy and MI for travel time deterministic are equal to 0 regardless of number of bins. However, for MI_{joint}, MI_{arith}, and MI_{geom}, MI for TIS is equal or close to 1.
Entropy performance compared to measures of error	The performance increases with the increase in the number of bins for all measures of errors. The best performance is found when compared with SAE (or SSE)	The three different optimum rules led to the same performance when compared with MAE or MSE and Sturges' rule led to the best performance when compared with SAE (or SSE)	The performance when compared with SAE or SSE increases with the increase in the number of bins and it is constant over the number of bins when compared with MAE or MSE	The performance when compared with SAE or SSE increases with the increase in the number of bins and it is constant over the number of bins when compared with MAE or MSE
MI performance compared to measures of dependence	In general, it showed better performance when compared with distance correlation, followed by Pearson correlation	In general, it showed better performance when compared with Pearson correlation, followed by $R^2{}_{adj}$	In general, it showed better performance when compared with Pearson correlation, followed by $R^2{}_{adj}$	In general, it showed better performance when compared with Pearson correlation, followed by $R^2{}_{adj}$
Comments	Number of bins appeared to be a significant factor for the performance of the method		Number of bins appeared to be a significant factor for the performance of the method	Normalization method did not appear to be a significant factor, however all normalization methods were analyzed together. If the methods were

					investigated only two by two, the results could have been different.

Table 54. Summary of results investigating information theory as a method for uncertainty quantification in simulation models using the kernel, KNN, fuzzy-histogram, and histogram-based methods, and Jayne's based approach as solution for the challenge encountered while applying entropy for continuous variables.

Method / Characteristics and results of the method	**Kernel with different values of bandwidth**	**Kernel with Silverman bandwidth**	**KNN**	**Fuzzy-histogram**	**Histogram-based with probability density function and fixed number of bins**
Characteristics of the method investigated	Kernel functions and bandwidth values	Kernel functions and Silverman bandwidth rule of thumb	Number of k-nearest neighbors	Fuzzy subsets and fuzzy membership functions	Number of bins
Number of bins or binwidth	NA	NA	NA	NA	Entropy and MI decrease with decrease in binwidth (increase in number of bins)
Bandwidth value	Entropy and MI increase with decrease in bandwidth	NA	NA	NA	NA
Number of k-nearest neighbors	NA	NA	Entropy and MI increase with decrease in the number of k-nearest neighbors	NA	NA
Number of fuzzy subsets	NA	NA	NA	Entropy and MI decrease with increase in the number of fuzzy subsets	NA
Number of replications	Entropy and MI increase with increase in number of replications for mid-range values of bandwidth.	Entropy and MI increase with increase in number of replications	Entropy and MI increase with increase in number of replications (for lower number of k-nearest	Entropy and MI increase with increase in number of replications (for lower number of fuzzy	Entropy and MI increase with increase in number of replications (for lower bins, the entropy and MI tend to decrease

	For low-values of bandwidth, the entropy and MI decreases with the increase in the number of replications. For high values, the entropy and MI are constant. The slope of the curve becomes shallower for any value of bandwidth with the increase in the replications.		neighbors, the curve tends to stabilize)	subsets, the entropy and MI tend to be constant)	with the increase in the number of replications)
Different seeds used	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"	The method resulted in same entropy for X_1 when using different seeds and different entropy and MI for other inputs and outputs. The method pointed larger variability in entropy of Y_1 and Y_2 when using "seed 3"
Different parameter values	Different entropy and MI when using different parameter values	Different entropy and MI when using different parameter values	Different entropy and MI when using different parameter values	Different entropy and MI when using different parameter values	Different entropy and MI when using different parameter values

Different traffic intensities	Not able to identify any pattern with changes in traffic intensity	Not able to identify any pattern with changes in traffic intensity	Not able to identify any pattern with changes in traffic intensity	Not able to identify any pattern with changes in traffic intensity	Not able to identify any pattern with changes in traffic intensity
CONWIP system	Constant entropy that tends to go to 0 as value of bandwidth increases and MI that slighltly decreases or increases for NIS. MI also tends to go to 0 as value of bandwidth increases	NA	Entropy of NIS equal to 0 and MI involving NIS is equal to the entropy of the input	Entropy of NIS equal to 0 and MI involving NIS is equal to the entropy of the input	Constant entropy and MI for NIS, which tend to go to 0 as number of bins increase
Travel time deterministic	Constant entropy that tends to go to 0 as value of bandwidth increases and MI that slighltly decreases or increases for travel time deterministic. MI also tends to go to 0 as value of bandwidth increases	NA	Entropy of travel time deterministic equal to 0 and MI involving travel time deterministic is equal to the entropy of the output	Entropy of travel time deterministic equal to 0 and MI involving travel time deterministic is equal to the entropy of the output	Constant entropy and MI for travel time detertministic, which tend to go to 0 as number of bins increase
Entropy performance compared to measures of error	The performance when compared with SAE or SSE increases with the increase in the value of bandwidth and it is constant over the value of bandwidth when	The two different kernel functions led to the same performance when compared with MAE or MSE and Epanechnikov function led to the best performance when	The performance when compared with SAE or SSE increases with the increase in the number of k-nearest neighbors and it is constant over the value of	The performance when compared with SAE or SSE increases with the increase in the number of fuzzy subsets and it is constant over the value of bandwidth	Not investigated

	compared with MAE or MSE	compared with SAE (or SSE)	bandwidth when compared with MAE or MSE	when compared with MAE or MSE	
MI performance compared to measures of dependence	In general, it showed better performance when compared with distance correlation	In general, it showed better performance when compared with distance correlation	In general, it showed better performance when compared with distance correlation	In general, it showed better performance when compared with distance correlation	Not investigated

3.5. Concluding remarks

This section was an extension of section 2 where Shannon's entropy and mutual information calculated using different estimators were investigated as potential measures to quantify uncertainty in simulation models. In section 2, histogram-based method was the only estimator considered. In this section, the following estimators were considered: kernel, KNN, and fuzzy-histogram.

In this section, the challenges encountered while applying entropy measures for continuous variables and the issues of interpretability faced when using the method proposed by Jaynes (1962) have been discussed more in depth. Based on the issues, an alternative for the function $m(x)$ was proposed, which was used to calculate the entropy and MI measures throughout this section of dissertation. The issue of calculating MI when using kernel estimator due to the multivariate interaction among bandwidths was also discussed. Similarly, a solution for this issue was proposed and applied when using the kernel method as an estimator for the entropy and MI measures.

In sections 3.4.3, 3.4.4, and 3.4.5, the impact of different bandwidths, different k-nearest neighbors, and different number of fuzzy subsets on the entropy and MI measures was discussed. In those sections, it was showed why the bandwidth in the kernel method had different impact on the entropy and MI measures than the binwidth in the histogram-based method. How this relates to the number of k-nearest neighbors and number of fuzzy subsets was also discussed.

In section 3.4.6, the impact of different traffic intensities, different seeds, different parameter values, and different systems on the entropy and MI measures was investigated. Some important observations from this section were that regardless of the method chosen and the number of bins used: (i) the entropy measure was able to correctly identify that X_1 have the same information/uncertainty among the different traffic intensity experiments; and (ii) the entropy measures indicate differences in information/uncertainty based on different seeds, different traffic intensities, and different parameter values. All the methods showed up as good alternatives to investigate the quality of a seed in simulation models. Nevertheless, the main observations of section 3.4.6 involved the CONWIP and deterministic travel time systems. Regarding those systems, all the methods were able to somehow capture the deterministic behavior of the input (or output). However, not all the methods were able to appropriately quantify the deterministic behavior or to capture the lack of impact of the input on the output (or vice-versa). While the KNN and the fuzzy-histogram methods were able to appropriately quantify the deterministic behavior of the input (or the output) through the entropy measure, the kernel method was only able to somehow capture the deterministic behavior. All three methods were not able to fully

capture the lack of impact of the input on the output through the MI. With that said, the entropy and MI measures appeared to have the best results in capturing the characteristics of the CONWIP or deterministic travel time systems when estimated using the histogram-based method with discrete empirical distribution. Using this latter method, both the entropy and the MI measures were able to fully capture and quantify the deterministic behavior of the CONWIP and the travel time systems.

In section 3.4.8, the results of the entropy measures were compared to SAE, SSE, MAE, and MSE, and the results of the MI measures were compared to distance correlation, Pearson correlation, and $R^2{}_{adj}$ to identify whether the measures agreed with other methods from the literature. Important conclusions from this section are: (i) the results of each method compared to SAE were always identical to the results compared to SSE, (ii) the results compared to MAE or MSE were constant over the different values of bandwidth, different number k-nearest neighbors, or different number of fuzzy subsets, and, (iii) all the methods showed better performance when compared to distance correlation, which was expected but different than what was found in section 2. It is important to highlight that the comparison with the different measures of error and different measures of dependence is not intended to give a rank on the best method to use or not. Rather, these comparisons are intended to give some information with respect to: (i) which factors/variables are more important for each method, (ii) a potential good choice of bandwidth, number of k-nearest neighbors, or number of fuzzy subsets for each method, and, (iii) a general idea of how the method performs in comparison with other methods available in the literature.

As the main limitations of this section, there are: (i) only a queue-system was used as an example to run the experiments; (ii) although the method was compared with other well-known measures of the scientific community, the method was not validated theoretically; and, (iii) a quantitative way to rank the methods among themselves was not developed. As future research, it is recommended: (i) to run similar experiments in different systems and to investigate how the responses would change (e.g., flow system, infection-transmission system, etc.), (ii) to propose a framework to validate the work theoretically, and, (iii) to apply the entropy and MI measures in a simulation model to estimate uncertainty and investigate its usability for simulation modelers.

Based on the results discussed, the recommendation while using the method is to calculate the entropy and MI measures using the histogram-based method with discrete empirical distribution, which although is not the correct estimator given the continuous nature of the variables, it was the estimator that showed the best results in detecting the uncertainty in the simulation models and it did not exhibit challenges for its estimation. If one decides to not follow this recommendation, the next recommendation is to follow the approach proposed by Jaynes (1962) but to use $m(x) = \hat{f}(x)\left(1 + \hat{f}(x)\right)$ as proposed in this section. Then, calculate entropy and MI measures using either the KNN or fuzzy-histogram, as these are the methods that showed the best quality overall results in quantifying uncertainty in simulation models.

4. APPLICATIONS OF ENTROPY MEASURES AS METHOD FOR INPUT PARAMETER SELECTION AND EXPERIMENT PLANNING IN SIMULATION MODELS

4.1. Introduction

The concept of entropy was coined in the physical sciences in the 19th century, more specifically in the area of thermodynamics (Greven, Keller, & Warnecke, 2014; Lacasa & Just, 2017). The term was introduced to describe energy dispersion, equilibria, and disorder of thermodynamic systems. Boltzmann later introduced the view of entropy as a measure of disorder of molecules in gas-related system. Thermodynamic systems share a number of common characteristics with complex systems and, consequently, have inspired the application of entropy theory for complex analysis (Mu & Hu, 2018).

In the past, real-world systems were thought to be accurately represented by linear cause-effect relationships. However, many of these real-world systems, such as epidemiological systems, the human immune system, the stock market, and aviation, exhibit complex dynamics that are difficult to quantify and cannot be represented by linear relations (Rickles, Hawe, & Shiell, 2007). The recognition of these complex systems has promoted research in developing measures and methods to quantify the uncertainty and complexity of many physical, biological, physiological, and socio-economic real-world systems (Xiong et al., 2017). Among the methods developed to quantify complexity, uncertainty, and the amount of information present in real-world systems, entropy measure

has become one of the most common methods (Amigo, Keller, & Unakafova, 2015; Lacasa & Just, 2017; Xiong et al., 2017).

In information theory, a branch of probability theory, entropy is also referred as information entropy to not be confused with thermodynamics entropy and it is interpreted as a measure of predictability, uncertainty, complexity, surprise, and information (or ignorance) (Tuan D. Pham, 2016; Rhea et al., 2011). Entropy as an information measurement method was first proposed by Shannon in a paper in 1948 (Shannon, 1948). This paper was later reprinted with corrections in 2001 (Shannon, 2001). According to Shannon (2001), the main problem of communication systems is to exactly or approximately reproduce at one point a message selected at another point considering that the actual message was selected from a set of possible messages. This problem can be reapplied to different contexts and fields.

Although information entropy measures have been initially applied as a measure of uncertainty and production rate of new information in the field of communication systems, over the past few decades many other information entropy measures have been proposed and applied in a wide range of fields (Xu, Ning, Chen, & Wang, 2004). According to Attaran and Zwick (1987) and Mousavian, Kavousi, and Masoudi-Nejad (2016), information theory has expanded to different research areas including: computer science, statistics, physics, management, marketing, finance, accounting, economics, neurobiology, bioinformatics, and systems biology. As examples of application fields, Xiong et al. (2017) cite: cerebrovascular dynamics, electroencephalography, heart rate

variability, financial time series analysis, gait and posture, earth sciences, imaging, cellular automata, feature classification, and others.

Despite its broad field of applications, the practical application of entropy measures is challenging, especially with respect to the performance evaluation of these measures and their correct interpretation. Among the main challenges are: (i) the variety of existing entropy measures and estimators, whose specificities are usually not completely understood by the users; (ii) the computation of information entropy is often non-trivial and the estimation approaches developed to approximate the probability density function utilized in the computation differ in their assumptions, which can make their performance assessment subjective; (iii) the difficulty in accurately estimating the probability function from the data without biasing the results; (iv) the difficulties in computing entropy measures for continuous variables; (v) the entropy measures can be affected by process-specific parameters and the complexity of the data, which can vary among different entropy measures and different estimators; (vi) the number of data points used in the calculation of entropy measures has been shown to affect the measures and for time-series data, the frequency at which the data is sampled is also relevant; (vii) the measures of entropy are also differently affected by noise; (viii) the entropy measures can be affected by how the data is partitioned and for entropy calculation, data usually has to be divided into bins or clusters and the number of possible bins approaches is large; (ix) the fact that many of the entropy measures proposed after Shannon's developments do not measure the same characteristics and they may violate some of the essential properties required from a measure of uncertainty; and, (x) the difficulty in comparing entropy

among different variables and among different studies because of large variability and lack of standards (Amigo et al., 2015; Dionisio et al., 2004; Estévez et al., 2009; Kapur, 1983; Kinney & Atwal, 2014; Mousavian et al., 2016; Rhea et al., 2011; Strehl & Ghosh, 2002; Tesmer & Estévez, 2004; Xiong et al., 2017). These challenges raises the following questions in the field of entropy: (i) is there a potential optimum partition (i.e., number of bins)?; and, (ii) is there a potential optimum number of data points to be used?

Rhea et al. (2011) highlighted that although challenging, entropy measures have shown to be a very useful tool for quantifying uncertainty and complexity. What is required is that researchers understand the underlying challenges and limitations of the measures and be aware of how the measures are affected by the different factors. Nevertheless, these challenges are likely some of the reasons why, to the best of our knowledge, entropy measures have not been extensively explored in the simulation field yet.

Simulation modeling is a suitable tool to investigate the dynamics of complex systems (Cicirelli, Furfaro, & Nigro, 2011; Hongqiao, Xihua, Fei, & Weizi, 2009; Venkatramanan et al., 2018; Xie et al., 2014a). For simulation models, a common goal is to answer "what-if" questions, i.e., to run a number of different scenarios and investigate how the response changes in each of them. The ultimate goal of any simulation modeler is to characterize the system performance to infer something about the real system and to optimize one or more of the system's responses (Dean & Lewis, 2006). In order to do so, simulation modelers need to first make assumptions about the real system and define the boundaries of the model. That is, simulation modelers must define what will be the scope

of the analysis and what will be left out. At this step, the input parameters that will be included in the simulation model and the responses or output parameters of interest are defined.

With the advances of technology and the increase of data availability, a challenge that arises is determining, among the many available data and input parameters, which are the most important to be included in the model (Ankenman & Nelson, 2012; Reshef et al., 2011). Adding data or parameters that are irrelevant and/or inaccurate to the simulation model is not only insignificant but most likely harmful for the results. This can increase the uncertainty of the model and lead to incorrect decision-making. There should be a balance between information loss and the computational resources needed (Haverkamp, Srinivasan, Frede, & Santhi, 2002). Even if the added data is correct, adding parameters to a model leads to increased resource needs. As highlighted by Oberkampf et al. (2002), a simpler model, with limited but known applicability, is more useful than a complete model with unknown applicability.

Due to real system abstraction and variability and in order to investigate different scenarios, running experiments is always necessary in studies using simulation models. In order to improve the use of computational resources and minimize the cost and time required for experimentation, simulation modelers must plan their computer experiments appropriately. This involves determining the run-length of the simulation model, the number of replications, and/or the number of scenarios to run and their configurations. A simulation model is inevitably uncertain and, therefore, a careful choice of the input parameters to be included in the model and the experimental plan can avoid waste of time

and resources and improve the accuracy of its results. For this, methods of uncertainty and information quantification, such as entropy measures, can prove useful.

In this section, Shannon's entropy and mutual information are proposed as measures of simulation uncertainty to support parameter selection and experiment planning in simulation models. Therefore, based on the context discussed, the main research question is: can entropy measures quantify the uncertainty present in simulation models and help to understand the input parameters and experiment settings, such as number of replications, number of bins, and seed, that contribute more to uncertainty in simulation models? In other words, can entropy and mutual information measures be applied to support the choice of input parameters and experiment settings for simulation models?

This section is intended to be a continuation of the previous sections of this dissertation.

The central contribution of this section of the dissertation is that although information theory has been widely recognized for its importance in the area of uncertainty and information quantification, the theory has not been extensively applied in the simulation field yet. In this work, the ability of entropy measures to quantify uncertainty in simulation models is investigated through a series of applications of the measures for input parameter selection and experiment planning. These applications involve calculating Shannon's entropy and mutual information for different queue simulation model experiments using stationary univariate distributions. The entropy measures are estimated using histogram-based method with probability density function and the entropy

normalization method proposed in section 3.4, as well as using histogram-based method with empirical discrete distribution. Regression analysis and ANOVA are also used to support the choice of experiment settings. Using Tukey-Kramer multiple comparison test and contingency analysis, the results of the entropy measures are compared to other methods such as standard error of the mean and ANOVA.

The rest of this work is organized as follows: section 4.2 provides a quick overview on some of the different existing entropy measures from the literature and different applications. Section 4.3 discusses the method used to investigate the potential of entropy measures to support the choice of input parameters and experiment settings. Results and analyses are reported in section 4.4. Concluding remarks and future research directions are presented in section 4.5.

4.2. Background

4.2.1. Entropy measures

Shannon defined entropy to be a statistical parameter that measures on average the information produced for each letter in a language (Shannon, 1951). Based on this, it is possible to investigate the predictability of a language when the preceding N letters are known. A representation of a general communication system is given in Figure 72, adapted from Shannon (2001) and Stone (2015). The information source issues a message (or sequence of messages). The transmitter encodes the message in some way to produce a signal suitable for transmission over the channel. The channel is the medium where the message is transmitted. The receiver decodes the message for the intended destination or

recipient. During the transmission process, the signal may be perturbed by noise. Hence, the received signal may not be exactly the same as the issued signal.

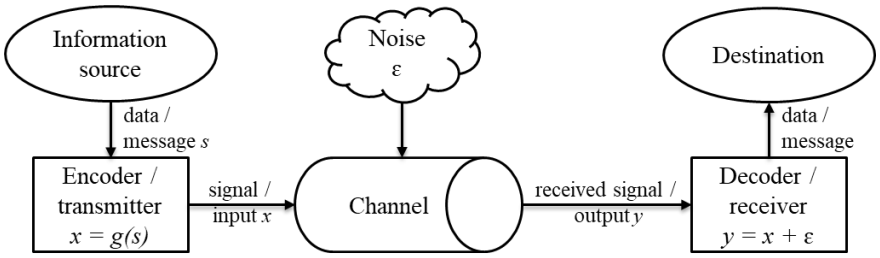

Figure 72. Representation of a general communication system.
Source: Adapted from Shannon (2001) and Stone (2015).

More generally, Shannon defined the information content of a random variable X as $h(x) = -\log p(x)$ and the entropy as the average information gained by knowing the outcome of the random variable X, which is equal to the average uncertainty removed (Amigo et al., 2015; Kapur, 1983; Mousavian et al., 2016; Xiong et al., 2017). For a discrete random variable, we have: $H(X) = -\sum p(x) \log p(x)$. According to this definition we have: (i) low information content for highly probable outcomes; and, (ii) high information content for unlikely outcomes (Xiong et al., 2017). Entropy is slightly more complicated to understand because it takes the average of the information content. Therefore, common outcomes contribute more to the entropy than rare outcomes, but they have less information content. In a succinct way, if there is no uncertainty, entropy is zero. On the other hand, entropy is maximum when all the outcomes are equally likely to occur (Telesca et al., 2008). He and Kolovos (2018) restated this in a more applicable approach:

the higher the probability that a model predicts an outcome, the less informative the model is. In other words, according to Shannon's entropy, a model that accounts for all possible outcomes and has no uncertainty, will predict the process correctly, yet it will provide no valuable information about the process itself. Figure 73 illustrates this idea by showing Shannon's entropy versus the bias of a coin. One can see that when the coin is fair ($p(heads) = p(tails) = 0.5$), the entropy is maximum and when the $p(heads) = 1$ or $p(heads) = 0$, the entropy is minimum.

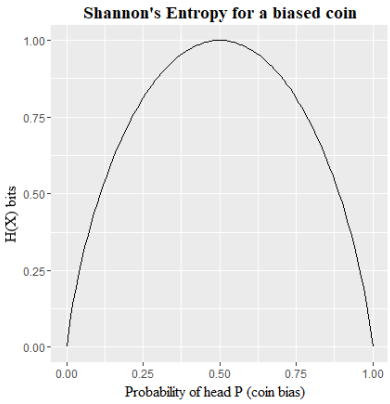

Figure 73. Shannon's entropy for a biased coin.
Source: Adapted from Stone (2015).

The Shannon's measure of information and uncertainty possesses many properties, such as: (i) symmetry - it should not change if $p_1, p_2, \dots, p_m$ are interchanged; (ii) continuity - it should be a continuous function of $p_1, p_2, \dots, p_m$; (iii) maximality - it should be maximum when all the probabilities are equal; (iv) additivity – it should be the sum of the entropies of two independent probability distributions; among others (Amigo et al.,

2015; Kapur, 1983). Without any formal mathematical derivation, Shannon proposed the entropy of a continuous random variable, called differential entropy, as: $H(X) = -\int p(x) \log p(x)\, dx$ (Xiong et al., 2017). The entropy of a discrete random variable X is a function of the distribution of the random variable and it depends only on the number of outcomes and the probabilities of the outcomes but not on the values of the outcomes taken by X (Mousavian et al., 2016; Stone, 2015). However, this is not true for the continuous random variable and its entropy depends on the range of values, which is one of the difficulties in defining entropy for continuous variables.

For the case of the noisy channel represented in Figure 72, Shannon (2001) listed a number of entropies that could be calculated:

- $H(X) = -\sum_X p(x) \log p(x)$: the entropy of the input of the channel or the average information per issued signal;
- $H(Y) = -\sum_Y p(y) \log p(y)$: the entropy of the output of the channel or the average information per received signal;
- $H(X,Y) = -\sum_X \sum_Y p(x,y) \log p(x,y)$: the joint entropy of input and output or the average information of the communication system associated to pairs of transmitted and received symbols;
- $H(X|Y) = -\sum_Y p(y) \sum_X p(x|y) \log p(x|y) = -\sum_Y \sum_X p(x,y) \log p(x|y)$: the conditional entropy of the input when the output is known or the average information measurement of the source given that Y was received. The entropy of X conditioned to the occurrence of a particular symbol y is given by: $H(X|y) = -\sum_X p(x|y) \log p(x|y)$; and,

- $H(Y|X) = -\sum_X \sum_Y p(x,y) \log p(y|x)$: the conditional entropy of the output when the input is known or the average information measurement of the received signal given that X was transmitted (i.e., the average uncertainty of Y that cannot be attributed to X or the noise entropy).

Shannon (2001) has also shown the following important inequality $H(X) \geq H(X|Y)$ and defined rate of transmission or, as it is more commonly called, mutual information. Mutual information is another important information measure and it measures how much of the entropy in the output reflects information in the input and how much is noise or what is the average reduction in uncertainty about the value of Y provided by the value of X and vice-versa (Mousavian et al., 2016; Stone, 2015; Vinh, Epps, & Bailey, 2010). In other words, MI measures the amount of information contained in a variable in order to predict the dependent one and it is a measure of the dependence and nonlinear relationship between the variables (Estévez et al., 2009; Fraser & Swinney, 1986; Kinney & Atwal, 2014; Rossi, Lendasse, François, Wertz, & Verleysen, 2006; Schreiber, 2000).

The MI $I(X;\ Y)$ can be calculated using Equation 73 (Dionisio et al., 2004).

$$I(X;Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y|X) = H(X) - H(X|Y) \quad \text{Equation 73}$$

For the noisy channel, we have:

$$Y = X + \varepsilon \quad \text{Equation 74}$$

If the value of X is known, the uncertainty in X is zero and, consequently, the entropy $H(X)$ is zero. Therefore:

$$I(X;Y) = H(Y) - H([X + \varepsilon]|X) \quad \text{Equation 75}$$

According to Stone (2015), because the uncertainty in X is zero, it makes no contribution to the conditional entropy, which yields:

$$I(X;Y) = H(Y) - H(\varepsilon|X) \quad \text{Equation 76}$$

But the value of the noise ε is independent of the value of X, which shows that $H(Y|X)$ is the noise entropy:

$$H(Y|X) = H(\varepsilon|X) = H(\varepsilon) \quad \text{Equation 77}$$

From Equation 73 and because $H(X) \geq H(X|Y)$, we know that $I(X;Y) \geq 0$, with equality only when X and Y are strictly independent (Dionisio et al., 2004; Kraskov et al., 2004). From Equation 73, one can also see that the MI of a random variable with itself as given by $I(X;X) = H(X) + H(X) - H(X,X) = H(X) - H(X|X) = H(X)$ is equal to the entropy of the variable and it is also known as self-mutual information (Mousavian et al., 2016).

The five aforementioned entropies and the MI can be calculated using knowledge of logarithm functions and probability theory, such as: marginal distribution, conditional distribution, joint distribution, and Bayes' rule. Using Bayes' rule, one can calculate the joint entropy by $H(X,Y) = H(X|Y) + H(Y) = H(Y|X) + H(X)$ and the marginal entropy by $H(X) = I(X;Y) + H(X|Y)$. Figure 74 shows the relationships between Shannon's entropy measures and MI, as well as the use of MI as a measure of dependence.

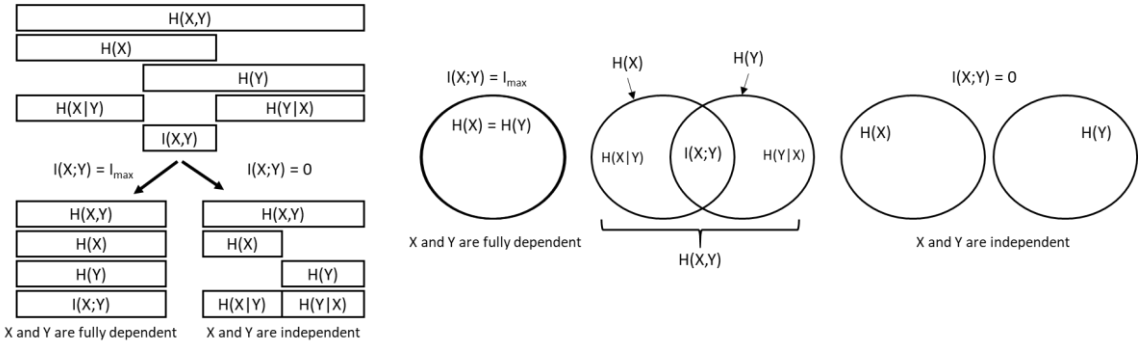

Figure 74. Representation of the relations between Shannon's entropy measures and mutual information and their application as a measure of dependence.
Source: Adapted from Bao-Gang and Yong (2008) [left-side], and Tesmer and Estévez (2004) and Stone (2015) [right-side].

From Figure 74, it is easy to see that MI is bounded below by 0 and bounded above by the minimum of the entropies yielding $0 \leq I(X;Y) \leq \min\left(H(X), H(Y)\right)$. Because the entropy of variables can vary significantly, a normalized version of MI is desirable for easier interpretation and comparisons (Strehl & Ghosh, 2002). Equation 78, proposed by McDaid, Greene, and Hurley (2011) and Vinh et al. (2010), provides a series of upper bound relations for MI.

$$I(Y;X) \leq min\left(H(X), H(Y)\right) \leq \sqrt{H(X)H(Y)} \leq \frac{1}{2}\big(H(X) + H(Y)\big) \leq max\left(H(X), H(Y)\right) \leq H(X,Y) \quad \text{Equation 78}$$

Several normalizations of MI are possible based on Equation 6. These normalizations include using the arithmetic or geometric mean of $H(X)$and $H(Y)$.

Different authors favor different normalizations. Strehl and Ghosh (2002) preferred to use the geometric mean as normalization because of the analogy with a

normalized inner product $NMI_{geo\ mean} = \frac{I(X;Y)}{\sqrt{H(X)H(Y)}} = \frac{I(X;Y)}{\sqrt{I(X;X)I(Y;Y)}}$. McDaid et al. (2011) argued that the most intuitive normalization would be to use $NMI_{max} = \frac{I(X;Y)}{\max(\max(H(X)),\max(H(Y)))}$. Principe et al. (2000) opted to use $NMI_{joint} = \frac{I(X;Y)}{H(X,Y)}$. NMI_{joint} is also known as information quality ratio (IQRt) (Wijaya et al., 2017). According to Wijaya et al. (2017), by considering the total uncertainty one has a ratio that gives a fairer comparison. Bao-Gang and Yong (2008) argued that the following normalizations $NMI_X = \frac{I(X;Y)}{H(X)}$ and $NMI_Y = \frac{I(X;Y)}{H(Y)}$ are not as rigorous because they produce unequal values due to the asymmetric property in their definition. So, the authors suggest the use of $NMI_{geo\ mean}$ or $NMI_{arith\ mean} = 2\frac{I(X;Y)}{H(X)+H(Y)}$. $NMI_{arith\ mean}$ was also suggested by Ghosh (2002). Estévez et al. (2009) proposed the use of $NMI_{min} = \frac{I(X;Y)}{\min(H(X),H(Y))}$. All these NMI are bounded in [0,1], where 0 means the two variables are independent and do not share any information about each other, and 1 means the two variables are identical and by knowing one variable, the next variable can be perfectly predicted (McDaid et al., 2011).

The mutual information index (MII) method was proposed by Critchfield and Willard (1986) based on the concept of mutual information. Critchfield and Willard (1986) used the MII to quantify the impact of uncertainty in the values chosen for the probabilities and utilities in a decision tree example. The MII provided insight about the contribution of the join effects of all tree variables to the decision tree under investigation. The measure can be calculated and compared for different inputs to determine which inputs provide

useful information about the output, but calculation of the MII suffers from computational complexity and the MII is easier computed for dichotomous responses (Frey & Patil, 2002). The MII is given by $MII_{X,Y} = \frac{I(X;Y)}{I(Y,Y)} \times 100\%$. $I(Y,Y)$ is the self-mutual information, which is equal to $H(Y)$. One can easily see that MII is just a normalization of the MI based on the self-mutual information of the output, as discussed in Bao-Gang and Yong (2008). According to Critchfield and Willard (1986), MII measures the percentage of the average mutual information contributing to the decision indicator variable Y that can be assigned to the tree variable X.

Reshef et al. (2011) presented the maximal information coefficient (MIC) as a measure of dependence for two-variable relationships. The idea is that if a relationship exists between two variables, either linear or non-linear, then the MIC can be calculated on the grid drawn on the scatterplot of the two variables. The MIC of the x-by-y grid applied to the data is given by $MIC_{X,Y} = \frac{I(X;Y)}{min\{H(X),H(Y)\}}$ (Z. Zhang, Sun, Yi, Wu, & Ding, 2015). All grids are explored up to a maximal grid resolution and, for every pair (x, y), the largest possible mutual information achievable is computed. MIC is between 0 and 1, symmetric, and robust to outliers (Reshef et al., 2011; Z. Zhang et al., 2015). Similar to the MII, MIC is just a normalization of the MI as discussed in Estévez et al. (2009). Kinney and Atwal (2014) highlighted that MI is a more natural and practical measure to equitably quantify associations in large datasets and that MIC is completely insensitive to noise.

After Shannon's developments, many other entropy-like quantity measures of information have been proposed (Wehrl, 1978). These other mathematical entropies do not measure the same characteristic and their definitions have been motivated by quite

different considerations (Kapur, 1983). According to Amigo et al. (2015) these measures have been developed as an attempt to generalize the axioms proposed by Shannon and due to the generalization, Kapur (1983) highlighted that they may violate some of the essential properties required (or expected) from a measure of uncertainty or information. Such generalized entropies, include information entropy measures introduced by Rényi (1961), Havrda and Charvát (1967), and Tsallis (1988).

Xiong et al. (2017) classified entropy measures into two groups: (i) static and (ii) dynamic. Static entropy measures do not take any temporal information into account when measuring an observed probability distribution and dynamic entropy measures study the information content of a stochastic process evolving in time. Shannon's entropy is a static measure that does not take any temporal information into account. Between 1950 and 1960, Kolmogorov and Sinai, formalized the concept of information theory for dynamic systems (Wehrl, 1978; Xiong et al., 2017). The Kolmogorov-Sinai entropy (KS), also known as conditional entropy, measure-theoretic entropy, or metric entropy, was developed to quantify the new information contained in the present but not in the past or the average rate of newly created information in dynamical systems through a sequence of observations (Amigo et al., 2015; Xiong et al., 2017). As a result, KS is widely used as a measure of randomness and predictability in dynamical systems (Amigo et al., 2015; Xiong et al., 2017). Katok (2007) provides a historical overview of the KS in the field of dynamics.

In practical terms, it is not possible to calculate the KS entropy for $N_t \to \infty$ because the measure diverges to infinity when the signal is contaminated by the slightest noise

(Orozco-Arroyave, Arias-Londono, Vargas-Bonilla, & Nöth, 2013). Therefore, Pincus (1991) proposed an estimation method for measuring the average conditional information generated by diverging points on a trajectory in state space. The method is called approximate entropy (ApEn). ApEn describes the production rate of new information or the repeatability and predictability of a time series by measuring the likelihood that a pattern with length (N) and criterion of similarity (r) at time delay (t) will repeat in the time series (Rhea et al., 2011; Yentes et al., 2013). As a result, frequent appearance of similar wave segments leads to lower values of ApEn. ApEn has the capability to identify changing complexity in quickly changing signals (Xu et al., 2004).

The ApEn depends on the signal length due to the self-comparison of points in the attractor (Orozco-Arroyave et al., 2013). That is, the ApEn algorithm counts each sequence as matching itself, which leads to bias (Bravi, Longtin, & Seely, 2011; Richman & Moorman, 2000). Sample entropy (SampEn) was designed by Richman and Moorman (2000) as an estimation method to overcome this limitation. SampEn, similar to ApEn, quantifies the change in the relative frequencies of length m time-delay vectors, but SampEn excludes the counts where the vector is compared with itself, which avoids the bias that self-matches introduce in the estimation (Amigo et al., 2015; Bravi et al., 2011). Besides eliminating self-matches, SampEn is also simpler to calculate than ApEn (Richman & Moorman, 2000).

Similar to Shannon's entropy, the original definition of mutual information does not contain dynamical nor directional information about the systems. Schreiber (2000) stated that one could incorporate dynamical structure in the mutual information by

introducing transition probabilities in place of static probabilities in the entropy and a somewhat ad hoc directional structure by introducing a time lag in one of the variables of the mutual information (Schreiber, 2000).

When the mutual information reflects the amount of information shared between the present and the past observations of a stochastic process, it is known as information storage (Xiong et al., 2017). The information storage measures the amount of information preserved in a time-evolving system or how much of the uncertainty about the present can be resolved by knowing the past. Similar to KS, the information storage can also be used to predict the future dynamics of the system. However, somewhat different than KS, information storage uses the past information for that and as a result, if the process is fully random, the past gives no knowledge about the present and the information storage is zero; on the contrary, if the process is fully predictable, the present can be fully predicted from the past, which results in maximum information storage. Finally, if the process is stationary, the information shared between the present and the past is constant (Xiong et al., 2017). In the context of dynamical systems, the entropy measures the information contained in the present state; the conditional entropy measures the new information that cannot be inferred from the past; and the information storage measures the information that can be explained by the past.

Abásolo, Escudero, Hornero, Gómez, and Espino (2008) presented the definition of auto-mutual information (AMI) and cross-mutual information (CMI). Auto-mutual information is the time-delayed self-mutual information of a signal and it quantifies, on average, the degree to which $x(t+\tau)$ can be predicted from $x(t)$. AMI is given by

$I(X;X)_t = \sum_{x(t),x(t+\tau)} P_{XX_t}[x(t), x(t+\tau)] \times \log\left\{\frac{P_{XX_t}[x(t),x(t+\tau)]}{P_X[x(t)]P_{X_t}[x(t+\tau)]}\right\}$. Cross-mutual information is the time-delayed mutual information of two different signals.

In the context of statistical inference, Kullback and Leibler (1951) proposed a measure of distance or divergence between two probability distributions, which measures the average number of extra bits needed to construct an optimal encoding if a different distribution q_i is used (Şahin, 2017; Schreiber, 2000). The Kullback-Leibler divergence (KL) is also known as relative entropy or information for discrimination (Mousavian et al., 2016).

The first term of the KL is called cross-entropy (Hopper, 2021). The KL calculates the relative entropy between two probability distribution and the cross-entropy calculates the total entropy between two probability distributions. Both KL and cross-entropy are non-symmetrical. Considering two distributions P and Q, where P represents the measured or "true" theoretical data and Q represents the approximated P data through a model, the KL divergence is the average difference of the number of bits required for encoding samples of P using a code optimized for Q. That is, the relative entropy of P with respect to Q. The cross-entropy is used to quantify the discrimination information between two probability distributions for a given random variable or set of events and it measures the average number of total bits needed to identify an object from a set of possibilities if a different coding scheme Q is used instead of the original source coding scheme P (Şahin, 2017). There are many situations where cross-entropy needs to be

measured but P is unknown. An estimate of cross-entropy can be calculated by: $H(P,Q) = -\sum_{i=1}^{N} \frac{1}{N} log_2 Q(X)$.

Minimizing the cross-entropy is known as principle of minimum cross-entropy (MinxEnt) and minimizing KL is known as principle of minimum discrimination. In summary, the principles state that, given a priori distribution Q, one has to choose the distribution P with the minimum cross-entropy (or KL) from the ones that satisfy all the constraints (X. Chen, Kar, & Ralescu, 2012; Kapur & Kesavan, 1990). Both methods can be used as loss functions in machine learning and classification models (Brownlee, 2019). When used as loss function for optimizing a classification predictive model, the entropy of the class label should be zero and, consequently, both cross-entropy and KL calculate the same quantity (Brownlee, 2019). Because of that, the terms KL and cross-entropy as well as principle of minimum cross-entropy and principle of minimum discrimination are sometimes used interchangeably, as they are in De Boer, Kroese, Mannor, and Rubinstein (2005) and Shore and Johnson (1981). It is worth pointing out that when Q is compared against a fixed reference distribution P, cross-entropy and KL are identical up to an additive constant.

Still in the context of statistical inference, Jaynes (1957) proposed the maximum entropy principle (MaxEnt) as a method for estimating probability distributions from data when one does not have complete knowledge (Nigam, Lafferty, & McCallum, 1999). The principle states that when one makes inferences based on partial information, one must use the probability distribution that has the maximum entropy subject to whatever is known (Jaynes, 1957). In other words, one wants to find a target probability distribution

of maximum entropy subject to a set of constraints that represents the incomplete knowledge about the target distribution (Phillips, Anderson, & Schapire, 2006). If nothing is known, the distribution should be as uniform as possible which gives the maximal entropy (Amigo et al., 2015; X. Chen et al., 2012; Nigam et al., 1999; Phillips, Dudík, & Schapire, 2004). By using a distribution with higher entropy, one is less constrained or has more choices, which means that the decision agrees with everything that is known, but carefully avoids assuming anything that is not known (Berger, Della Pietra, & Della Pietra, 1996; Phillips et al., 2006). If Q, in MinxEnt is the uniform distribution, the principle reduces to the maximum entropy principle (Kapur & Kesavan, 1990). Berger et al. (1996) highlighted that in simple cases one could find the solution to the MaxEnt constraint programming analytically, but in general a more direct approach is needed.

MaxEnt and MinxEnt involve finding the maximum (or minimum) entropy when the constraints and a priori probability distribution are given. Kapur and Kesavan (1990) discussed the inverse problem, when a probability distribution is given and one has to find either: (i) the constraints, or (ii) the measure of entropy, or (iii) the priori probability distribution, so that the given probability distribution is a MaxEnt or MinxEnt.

Another entropy approach worth discussing is the Bayesian maximum entropy (BME). In the BME approach, two types of data are differentiated: hard data, whose observations are considered to be deterministic, and soft data, whose observations represent observed data that carry uncertainty. Instead of a single measure of information, BME is actually a knowledge-centered approach to integrate information from different sources to obtain improved prediction. BME is a model-free approach that combines

maximum entropy theory with Bayesian statistics to handle different types of information and uncertain input (He & Kolovos, 2018). The BME is applied in three phases: meta-prior stage, prior stage, and posterior stage. The advantage of BME is that it can be applied for spatial and spatiotemporal information, as well as univariate and multivariate cases.

The BME approach is based on Bayes' theory and the maximum entropy principle. Therefore, the BME approach says that based on a set of trial or prior probability distributions the prior distribution with the maximum entropy should be chosen. This prior is then used to update the current information and obtain the posterior distribution, following the Bayesian scheme. A little counter-intuitive is the fact that based on BME, the higher the probability that a model predicts a process, the less informative the model. The logic is that a model that predicts a process correctly, provides no valuable information about the process itself because no behavior is implied to explain the process (He & Kolovos, 2018). The measure of information commonly used in the context of Bayesian statistics is the KL of the posterior from the prior (Walsh, Wildey, & Jakeman, 2018).

In 1956, Lindley (1956) introduced the concept of the measure of the information provided by an experiment. Lindley's measure was derived from Shannon (1948) and expressed the knowledge prior to performing the experiment in terms of the prior probability distribution. In Shannon's work, by considering the information in x and y one can measure the rate of transmission of information along the channel. Hence, Lindley (1956) proposed an analogous description in the context of experimental design, where x represents the knowledge of the system prior to the experiment (prior), and y is the

knowledge after the experiment (posterior). The comparison of the knowledge before and after the experiment yields the amount of information provided by the experiment. Lindley's experimentation approach was to perform the experiment for which the expected gain in information was the largest and to continue this process until the desired amount of information about the system was achieved. Lindley (1956) used the measure to solve some experimental design problems where the goal was to gain knowledge about the system.

When one applies the Bayesian approach to find an experimental design that is optimum according to the modeler's goals, such as parameter inference, prediction, or model selection, the approach is known as optimal experimental design (Clyde, 2001). Optimal experimental designs aim at maximizing the value of each experiment and minimizing the uncertainties to achieve the experimental goals more rapidly and with lower costs (Elizabeth G. Ryan et al., 2014; van Den Berg, Curtis, & Trampert, 2003; Walsh et al., 2018).

The first studies on Bayesian optimal experimental designs followed the work of Lindley (1956) and suggested using the expected gain in information given by an experiment as a utility function (Chaloner & Verdinelli, 1995). The goal was then to maximize this utility function. According to Huan and Marzouk (2014) and Bisetti, Kim, Knio, Long, and Tempone (2016), a useful utility function for Bayesian parameter inference is the KL from posterior to prior. Huan and Marzouk (2014) highlighted that other utility functions may be used depending on the goal of the experiment, but the KL is a general-purpose function to maximize understanding about the system and,

consequently, it yields good results in a variety of experiments. For other utility functions, please refer to Chaloner and Verdinelli (1995) and Elizabeth G Ryan (2014).

Although useful, the computations of the KL utility function and the integration problem (posterior calculation) are not trivial (Bisetti et al., 2016; Huan & Marzouk, 2014). Unless the likelihood and the priors are appropriately chosen, such as the use of conjugate priors, or the problem is restricted to special cases, such as linear Gaussian models, usually one cannot find a closed form solution to the problem (Elizabeth G Ryan, 2014; Walsh et al., 2018). Therefore, the functions must be solved using numerical approximation methods or stochastic solution methods. Elizabeth G Ryan (2014) also highlighted the need to check the solution for sensitivity to the prior distribution.

In the context of feature selection in images, mutual information is used to measure the level of similarity or redundancy between pixels and the relevance of the features for the task classification. At least four mutual information measures were cited by Tapia and Perez (2013) and Estévez et al. (2009) for feature selection: (i) minimum redundancy and maximal relevance (mRMR), (ii) normalized mutual information feature selection (NMIFS), (iii) conditional mutual information feature selection (CMIFS), and (iv) conditional mutual information maximization (CMIM). mRMR combines relevance and redundancy into a single criterion: $f^{mRMR}(X_i) = I(C; f_i) - \frac{1}{|S|}\sum_{f_i \in S} I(f_i; f_s)$, where $I(C; f_i)$ measures the relevance of the feature to be added for the output class and $\frac{1}{|S|} I(f_i; f_s)$ estimates the redundancy of the $f_{i^{th}}$ feature with respect to the subset of previously selected features S. NMIFS is a normalized version of mRMR where the

mutual information $I(f_i; f_s)$ is normalized by the minimum entropy of both features. Thus, we have: $f^{NMIFS}(X_i) = I(C; f_i) - \frac{1}{|S|}\sum_{f_i \in S} \frac{I(f_i;f_s)}{\min\,(H(f_i),H(f_s))}$. In CMIFS, the subset of features S is built step by step by adding one feature at a time. Let S be the set of already-selected features and N the set of candidate features, such as $S \cap N = \emptyset$. The next feature in N to be selected is the one that maximizes: $I(C; f_i) - [I(f_i; X_s) - I(f_i; X_s|C)]$, where $f_i \in N$. Finally, CMIM considers the MI between the candidate feature variable f_i and the task classification class C given each of the variables in the set S separately, this allows CMIM to consider a feature f_i relevant only if it provides large amount of information about the class C and this information is not contained in any of the features already selected. For this, we have: $CMIM = \begin{cases} argmax_{f_i \in F}\{I(f_i; C)\}\ for\ S = \emptyset \\ argmax_{f_i \in F/S}\{min_{f_i \in S}\, I(f_i; C/f_j)\}\ for\ S \neq \emptyset \end{cases}$,

where F is the initial set of n features for the empty set S.

In the context of model selection, the Akaike information criterion (AIC) was designed as an estimate of the expected Kullback-Leibler divergence between the model generating the data and a fitted candidate model (Cavanaugh, 1997). The KL is just one kind of loss function. The AIC is an estimate of the quality of the model and, consequently, it is used as a regression model selection criterion. Because the true model is unknown, the absolute divergence between a candidate model and the true model is also unknown, but the relative differences between models can be used to rank order models according to their expected KL. The candidate model with the lowest AIC has also the lowest expected KL, even though the actual KL is unknown. In other words, the AIC tells nothing about the quality of the model, but only about the quality relative to other models. If all

models fit poorly, AIC will not capture that. This explains why the KL is explicitly defined for two models and the AIC only for one and why the AIC is only a consistent estimator of KL if the true model is among the models under consideration (Vrieze, 2012). The AIC asymptotically selects the model that minimizes the mean squared error of prediction (Vrieze, 2012). If the number of data points is small, some correction is usually necessary (Cavanaugh, 1997). Wand (1997) classified AIC, MinxEnt, and MaxEnt in a category called information-based criteria.

Some of these proposed entropy and mutual information measures discussed above and a few other ones are summarized in Table 55. However, many information measures have been mentioned in the literature, such as: (ε, τ)-entropy (Amigo et al., 2015); Kapur's entropy (Kapur, 1983); fuzzy cross-entropy (Şahin, 2017); recurrence quantification analysis entropy (Rhea et al., 2011); corrected conditional entropy, fuzzy entropy, compression entropy, permutation entropy, distribution entropy, and multiscale entropy (Bravi et al., 2011; Xiong et al., 2017); diffusion entropy (Bravi et al., 2011); spectral entropy (Garner & Ling, 2014); and, compression entropy (Baumert et al., 2004; Bravi et al., 2011; Truebner et al., 2006; Xiong et al., 2017). For more discussion on most of the measures discussed in this section, please refer to Kapur (1983) and Bravi et al. (2011).

Table 55. Some entropy and mutual information measures proposed in the literature.

Entropy measure	Definition	Characteristics	Source

Renyi's entropy	$H_{\propto}(X) = \frac{1}{1-\propto}\log\sum_{x\in M} p(x)^{\propto}$, where $\propto \geq 0$, $\propto \neq 1$ is the order of the entropy.	Generalize Shannon's entropy for the case of generalized distributions, where $\sum_{x\in M} p(x) \leq 1$. $H_0(X) = \log\lvert\Gamma\rvert$ is called the Hartley entropy and $H_2(X) = -\log\sum_{x\in M} p(x)^2$ is called the collision or quadratic entropy.	Rényi (1961), Kapur (1983), Principe et al. (2000), Kannathal, Choo, Acharya, and Sadasivan (2005), and Amigo et al. (2015)
Havrda and Charvat's entropy	$H^{\propto}(P) = \frac{1}{1-e^{1-\propto}}[1 - \sum_{i=1}^{m} p_i^{\propto}]$, where $\propto \neq 1$ is the order of the entropy.	Replaced the recursive property of Shannon's entropy. This measure is not additive.	Kapur (1983)
Tsallis' entropy	$H_{\propto}(P) = \frac{k}{\propto-1}(1 - \sum_{i=1}^{m} p_i^{\propto})$, where $\propto \neq 1$ is the order of the entropy.	This entropy differs from Havrda and Charvat's entropy only in a factor that depends on $\propto$.	Tsallis (1988) and Amigo et al. (2015)
Kolmogorov-Sinai entropy	$H_{KS} = \lim_{\tau\to 0}\lim_{\varepsilon\to 0}\lim_{N_t\to\infty}\frac{1}{N_t\tau}\left(H_{N_t} - H_0\right)$ is the average K-S entropy, where H_{N_t} is the Shannon entropy calculated at time N_t given by $H_{N_t} = -k\sum_{i=1}^{N_t} p_i \log p_i$, N_t is the number of time steps, τ is the length of time, and ε is the bin or cell size. For discrete systems, τ is set to 1 and the $\lim_{\tau\to 0}$ is dropped out.	The KS is also a measure of the information needed to predict which part of the space the dynamics will visit at a time $t+1$, given the trajectories up to time t. H_{KS} is zero for a deterministic system because it does not produce any new information, some positive constant for a chaotic or stationary system because the system produces new information at a constant rate, and infinite for a fully random process as it produces information at the maximum rate. KS is based on three key characteristics: (i) modeling of sequential probabilities, (ii) entropy rate, and (iii) limiting conditions.	Bravi et al. (2011), Orozco-Arroyave et al. (2013), Tuan D Pham (2013), Tuan D. Pham (2016), and Xiong et al. (2017)

Approximate entropy	$A_E(m,r,N) = \lim_{N\to\infty}[\Phi^{m+1}(r) - \Phi^m(r)] = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\log C_i^m(r) - \frac{1}{N-m}\sum_{i=1}^{N-m}\log C_i^{m+1}(r)$, where $\Phi^m(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\log C_i^m(r)$, where $C_i^m(r)$ is the correlation dimension: $C_i^m(r) = \frac{2}{N(N-1)}\sum_{j=i+1}^{N}\theta(r - \|x_i - x_j\|)$, where N is the number of points in the state space, θ is the Heaviside function and $\|.\|$ is a norm.	Given N points, the family of statistics $A_E(m,r,N)$ is approximately equal to the negative average logarithm of the conditional probability that two sequences that are similar for m points remain similar within a tolerance r at the next point. Consequently, a low value of ApEn reflects a high degree of regularity or less complexity. N is the data size, r is the criterion of similarity, and m is the length of the data segment being compared.	Pincus (1991), Richman and Moorman (2000), Xu et al. (2004), Kannathal et al. (2005), (Rhea et al., 2011), Orozco-Arroyave et al. (2013), and Yentes et al. (2013)
Sample entropy	$S_E(m,r) = \lim_{N\to\infty}\left[-\log\frac{\Gamma^{m+1}(r)}{\Gamma^m(r)}\right]$.	$S_E(m,r)$ is precisely the negative logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in calculating the probability. Therefore, a lower value of SampEn also indicates more self-similarity in the time series.	Richman and Moorman (2000), Bravi et al. (2011), Orozco-Arroyave et al. (2013), and Amigo et al. (2015)
Approximate entropy with Gaussian kernel	The Heaviside function in the ApEn is replaced by: $dG(x_i, x_j) = exp\left(\frac{-\|x_i - x_j\|_1}{10r^2}\right)$.	The Gaussian kernel is used to give greater weight to nearby points by replacing the Heaviside function.	Orozco-Arroyave et al. (2013)
Mode entropy (ModEn)	$ModEn(m,r,N) = \Phi^{m-1}(r) - \Phi^m(r)$, for $m \geq 1$, m and r are commonly referred to as embedding dimension and radius, respectively.	Different than ApEn that compares the element of the first vector with that of the second vector, ModEn compares the increment of each element of the first vector with that of the second vector in determining whether two vectors are approximate. With this method the case where approximate vectors are too few would not occur even if the fluctuation of wave were slow and large, which makes it more adequate for time series of short-term signals with broad amplitude and slow fluctuation.	Xu et al. (2004)

Information storage	$S(X) = I(X_n, X_n^-) = E\left[\log \frac{p(x_1,\ldots,x_n)}{p(x_1,\ldots,x_{n-1})p(x_n)}\right]$. According to Xiong et al. (2017), Shannon's entropy, K-S entropy, and information storage are related to each other by: $S(X) = H(X_n) - C(X_n)$, where $C(X_n)$ is the conditional entropy $H(X_n\|X_n^-)$.	It quantifies the amount of information carried by the present that can be explained by the past. In other words, it reveals the degree to which information is preserved in a time-evolving system. If the system is fully predictable, the present can be predicted by the past and the information storage is the maximum; if the system is fully random, the past gives no information about the present and the information storage is zero; if the system is stationary, the information shared between the present and the past is constant.	Xiong et al. (2017)
Transfer entropy	The transfer entropy is given by $T_{J\to I} = \sum \log \frac{p\left(i_{n+1}\middle\|i_n^{(k)}, j_n^{(l)}\right)}{p\left(i_{n+1}\middle\|i_n^{(k)}\right)}$.	Transfer entropy was introduced by Schreiber (2000) to measure the rate of information exchange between two signals in both directions separately in the context of time series analysis. By conditioning on transitioning probabilities, the transfer entropy is able to appropriately exclude exchanged information due to common history and input signals when desirable.	Schreiber (2000)
Directional entropy	The n-dimensional entropy function is given by $h_{\mu,n}(T,V) = sup_\alpha\, h_{\mu,n}(T,V,\alpha)$. Where V is an n-dimensional subspace of $\mathbb{R}^d$, $1 \le n < d$, and T a measure-preserving $\mathbb{Z}^d$ action on a Lebesgue probability space (Ω, β, μ), and α a finite partition of Ω.	Directional entropy was introduced by Milnor (1988) as a n-dimensional entropy function that measures the density of information in very large finite sets to describe the distribution and flow of information throughout $(n+1)$-dimensional lattice in the context of cellular automata. A directional version of mutual information was introduced by Massey (1990) to capture the directed information flow from a random sequence X^n to a random sequence Y^n. The directional version of mutual information is given by $I\left(X^n \to Y^n\right) = \sum_{n=1}^{N} I(X^n; Y^n\|Y^{n-1})$.	Milnor (1988)

Kullback-Leibler divergence	The Kullback-Leibler entropy of two distributions P and Q defined on the same probability space i is given by $D_{KL}(P\|\|Q) = \sum_i p_i \log \frac{p_i}{q_i} = \sum_i -p_i \log q_i + p_i \log p_i = \sum_i p_i \log \frac{1}{q_i} - H(P)$.	KL is always non-negative and is zero if, and only if, $p = q$ According to Schreiber (2000), when one assumes two processes X and Y are independent, the KL gives the mutual information $I(X;Y) = D_{KL}(P(X,Y)\|\|P(X)P(Y)) = \sum_{x,y} p(x,y) \log \frac{p(x)}{p(x)p(y)}$. Because the Kullback-Leibler divergence is difficult to estimate nonparametrically in high dimensional spaces, Principe et al. (2000) have proposed two approximations to it based on quadratic distances: the first is based on Cauchy-Schwartz inequality and the second is based on Euclidean distance.	Schreiber (2000) and Mousavian et al. (2016)
Cross-entropy	The first term of the KL divergence is the cross entropy between two distributions P and Q: $H(P;Q) = \sum_i p_i \log \frac{1}{q_i}$.	The cross entropy is an efficient method for rare event simulation and for solving NP-hard optimization problems, such as the travelling salesman problem (Rubinstein, 2001). The method was motivated by an adaptive algorithm for estimating rare event probabilities in complex stochastic network, which involves variance minimization (De Boer et al., 2005).	Hopper (2021)
Cressi-Read measure (CR)	The CR is given by $D_{CR}(P;Q;\gamma) = \frac{1}{\gamma(\gamma+1)} \sum_i p_i \left[\left(\frac{p_i}{q_i}\right)^{\gamma} - 1\right]$, where γ is a parameter that indexes members of the CR family, q_i are the reference probabilities, and p_i are the subject probabilities.	The Cressie–Read is an extension of KL divergence and it provides an objective assessment of how much information a given probability distribution contains relative to a second. Over defined ranges of the divergence measures, the CR and Renyi's and Tsallis' entropies are equivalent.	Kowalski, Martin, Plastino, and Judge (2012)
Jensen-Shannon divergence (JS)	The JS is given by $D_{JS}(P\|\|Q) = \frac{1}{2} D_{KL}(P\|\|M) + \frac{1}{2} D_{KL}(Q\|\|M)$, where $M = \frac{1}{2}(P+Q)$.	The Jensen-Shannon divergence is a smoothed and symmetrical version of the KL and it provides another way to quantify the difference between two probability distributions. Because JS is symmetrical, $D_{JS}(P\|\|Q) = D_{JS}(Q\|\|P)$.	Fuglede and Topsoe (2004)

4.2.2. Applications of entropy measures

According to Amigo et al. (2015), entropy is a concept that appears in different contexts with different meanings. Entropy as a measure of information, uncertainty, surprise, compression, disorder, irregularity, complexity, or homogeneity has been applied in many different fields (Attaran & Zwick, 1987; Lotfi & Fallahnejad, 2010). Although information entropy was initially developed to address problems of data storage and data transmission in the field of communications, in recent decades the practical applications of entropy in fields such as marketing, management, finance, accounting, mathematics, natural sciences, social sciences, physics, computer science, chemistry, biology, bioinformatics, economics, behavioral sciences, and geophysics have become unparalleled (Amigo et al., 2015; Attaran & Zwick, 1987; Mousavian et al., 2016; Xiong et al., 2017).

In the marketing field, entropy has been used as a measure of uncertainty to represent consumers' preferences for brands (Attaran & Zwick, 1987).

In management, entropy has been applied as a measure of industry competitiveness by comparing industrial diversity either among regions or for a particular region over time (Attaran & Zwick, 1987). Entropy was also applied to identify the geographical allocation of different types of industries through hypothesis testing (Attaran & Zwick, 1987).

In the financial context, entropy has been applied for portfolio selection in two different ways: (i) as a measure of the uncertainty of portfolio returns by applying the MaxEnt to find the most unbiased probability distribution of future security returns for investors given limited information (Huang, 2012); and (ii) as a measure of securities

portfolio risk, whose components yield stochastic returns, by replacing the traditional mean-variance models with the divergence measure of portfolio asset diversification (Attaran & Zwick, 1987; Huang, 2012; Qin, Li, & Ji, 2009). W. Zhang and Wang (2017) used KS and other descriptive statistics to analyze the statistical behaviors of an agent-based financial price model. The authors applied KS to understand the complexity properties of the proposed model that attempted to reproduce the nonlinear behavior of financial markets. Within the financial field, mutual information has been applied as a measure of dependence in financial time series, such as stock market indexes (Dionisio et al., 2004). The authors used MI to show that the rate of returns of the financial market were not independent and identically distributed.

In the accounting field, entropy has been used as a measure of the loss of information from aggregating items on financial reports, such as the balance sheet (Attaran & Zwick, 1987).

In urban planning, more specifically in water resources planning, entropy was used as a measure of the loss of information from aggregating spatial data, either through the number of subwatersheds or the number of raingauges, in complex hydrologic models designed to simulate larger watersheds or the total runoff in a watershed (Haverkamp et al., 2002; Hernandez, Lane, Stone, Martinez, & Kidwell, 1997).

In conservation biology, Phillips et al. (2004) used the MaxEnt, which turned out to be the same as minimizing the KL, to investigate the geographic distributions of bird species, a critical problem in the field. In their work, the environmental variables were the given set of features, the occurrence localities served as sample points, and the

geographical regions were the space on which the distribution was defined. Phillips et al. (2006) used the MaxEnt for modeling and prediction of the geographic distribution of a species of sloth and a species of a small montane rodent.

In the context of food processing and quality, Wijaya et al. (2017) used information quality ratio to determine the best mother wavelet for electronic noise that generates signals from gas sensors for beef quality classification. The measure led to the selection of the wavelet that was able to better reconstruct the noise signal by keeping essential information from the original signal and reducing the noise level.

In the context of experimental design, Malakar and Knuth (2011) presented an entropy-based search algorithm for efficient experimental design to select the most informative experiment from a set of potential experiments described by many parameters. According to the authors, the algorithm was capable to select the highly relevant experiments and it was also more efficient than brutal force search.

In the statistical inference context, entropy measures have been used to test for stochastic independence and determinism of time series, for detecting association and discrimination between random variables and distributions, for assessing the probability distribution estimation from time series data, among other applications. Amigo et al. (2015) used permutation entropy to test the null hypothesis that the noisy data were outcomes of an independent and identically distributed random process and the alternative was that the noisy data were outcomes of determinism. Reshef et al. (2011) used normalized mutual information to detect greatest associations between pairs of variables in large data sets. Şahin (2017) proposed a cross-entropy measure using interval

neutrosophic information to determine the information measure for discrimination between two interval neutrosophic sets in the field of fuzzy sets. Kowalski et al. (2012) used CR divergence measures to assess the probability density function estimation from time series data using two different methodologies, namely histogram and Bandt and Pompe's methodologies. The authors confirmed the superiority of Bandt and Pompe's methodology. Approximate entropy and sample entropy were used to measure the complexity of finite time series and the underlying dynamical systems by quantifying the change in relative frequencies of length k time-delay vectors (Amigo et al., 2015).

In the context of time series analysis, entropy is mainly used to quantify the complexity of both data and systems. Shannon's entropy, approximate entropy, sample entropy, permutation entropy, and other generalized entropies, such as Tsallis' entropy and Renyi's entropy, have been applied to physiological data analysis (Amigo et al., 2015).

For cardio-physiological data, entropy measures have been applied to investigate fetal abnormal heart-rate, to discriminate between healthy patients and patients with congestive heart failure, to evaluate heart rate in patients with dilated cardiomyopathy, and to detect ventricular tachycardia by analyzing heart-rate variability from electrocardiograms (ECGs) (Amigo et al., 2015; Bravi et al., 2011). Xu et al. (2004) introduced ModEn to identify myocardium infarction in the high frequency electrocardiogram data of an animal model. Silva, Silva Filho, Crescêncio, and Gallo (2012) used KS, autoregressive integrated moving-average model (ARIMA), and the largest Lyapunov exponents to identify the anaerobic threshold in the heart rate time series

of a group of healthy men in rest and dynamic exercise in a seated position. According to these authors, physical exercise is a complex system involving physiological processes, movements of body segments, and other processes not well-known. The authors concluded that there was a strong correlation among the three methods in detecting the anaerobic threshold. Xiong et al. (2017) investigated the ability of entropy, conditional entropy, and information storage in detecting changes in the static and dynamical properties of heart rate variability in individual in different physiological states (awake and sleepy state) and clinical conditions (healthy and congestive heart failure subject). The entropy measures were calculated using linear, kernel, and nearest-neighbor estimators and the analyses were performed under three types of data preprocessing procedure. According to these authors, the entropy measures could only identify changes in cardiac dynamics in specific cases. Next, the authors recommended which measures to use depending on the purpose of the analysis and highlighted the need for appropriate data preprocessing and careful interpretation of the results based on the properties of the specific chosen entropy measure and estimator.

For brain-physiological data, entropy measures have been applied to distinguish between healthy patients and patients with Alzheimer's disease, to detect epileptic seizures, to identify sleep stages, and to quantify the effects of anesthetic drugs on brain activity by analyzing brain activity from electroencephalograms (EEGs) and/or magnetoencephalograms (MEGs) (Abásolo et al., 2008; Amigo et al., 2015). Kannathal et al. (2005) used Shannon's entropy, Renyi's entropy, Kolmogorov-Sinai entropy and approximate entropy to investigate normal and epileptic signals in EEGs. Z. Zhang et al.

(2015) used maximal information coefficient to measure brain functional connectivity. According to the authors, when compared with other measures of brain functional connectivity, such as correlation coefficient and coherence, MIC performed the best in terms of consistency and robustness.

Other physiological studies involving entropy measures include the investigation of gait and postural complexity in individuals and the use of speech signals. Yentes et al. (2013) explored the effects of changing the input parameters m, r, and N on the robustness of ApEn and SampEn. According to the authors, the results were very sensitive to the parameters, especially for small data sets. They applied the methods to investigate gait in young and older adults. Rhea et al. (2011) studied how noise, sampling frequency, and time series length affected approximate entropy, sample entropy, and recurrence quantification analysis entropy measures when the measures were applied to human center of pressure data as a measure of human postural control complexity. Orozco-Arroyave et al. (2013) applied four different entropy measures, namely approximate entropy, approximate entropy with Gaussian kernel, sample entropy, and sample entropy with Gaussian kernel, to detect Parkinson's disease in people through the use of speech signals. The authors also used six other nonlinear dynamic measures, such as Lyapunov exponent and Hurst exponent, to classify speech signals of people with Parkinson's disease from the control set.

In systems biology, entropy was used to find subnetwork markers in classification of cancer samples and determining optimal gradient sensing strategies in chemotaxing cells (Mousavian et al., 2016). Compression entropy was used by Baumert et al. (2004)

and Truebner et al. (2006) for forecasting life-threatening tachycardia in patients. Entropy measures have also been used to estimate complexity in hormonal patterns, blood pressure, and human postural control as a result of aging, disease, and/or disorder (Rhea et al., 2011).

In physics, Blanco, Casini, Hung, and Myers (2013) discussed potential uses of relative entropy between two states for vacuum state tomography.

In machine learning, entropy has been used for training, redundancy reduction, and prediction through entropy minimization, and mutual information has been used for independent component analysis, blind source separation, feature extraction, feature selection, classification, and information filtering (Principe et al., 2000). For an explanation of the difference between feature extraction and feature selection, please refer to Estévez et al. (2009). Wang and Hu (2009) added that information entropy measures have also been applied to model evaluation directly. Principe et al. (2000) used Renyi's quadratic entropy and two approximations to KL based on quadratic distances to train linear and nonlinear mappers in the context of unsupervised and supervised machine learning. The authors presented an algorithm to train learning machines to maximize (or minimize) mutual information between their input and output, which can be estimated through KL. The authors showed two different applications of the algorithm: one for feature extraction for classification (supervised learning) of vehicles in synthetic aperture radar imagery and the other for blind source separation (unsupervised learning), which is a linear mixture of independent source signals from which no further information about their sources is available.

In a similar way, Wang and Hu (2009) used normalized mutual information in the field of pattern recognition to assess the performance of classifiers. Instead of using the measure for feature extraction, the authors applied it for model evaluation directly. For that, the authors used the maximum normalized information criterion.

Maximum entropy principle was used in various natural language tasks, such as text segmentation, text classification, part-of-speech tagging, and language modeling (Nigam et al., 1999). Berger et al. (1996) presented an approach consisting of two tasks for natural language prediction using the principle of maximum entropy. The first task involved determining a set of features that captures the behavior of the random process (e.g., speech translation), and the second task involved selecting the most accurate model that encompasses these features to predict the output of the process (e.g., word pronunciation).

Tapia and Perez (2013) investigated four different mutual information measures, namely mRMR, NMIFS, CMIFS, and CMIM, in the context of feature selection to improve gender classification of face images.

Tuan D Pham (2013) used KS to investigate the spatial content of images and as a multidimensional feature for pattern classification. According to the author, the method was effective in detecting spatial characteristics of different scenes and spatial objects. Tuan D. Pham (2016) used KS entropy in the setting of image texture analysis and classification by quantifying uncertainty of pixel distributions in images.

Within feature selection algorithms, cluster analysis is needed in order to find the number of clusters that is associated with the optimum number of features (Wijaya et al.,

2017). In the context of cluster analysis, entropy has been used as a measure of diversity or dissimilarity (Amigo et al., 2015). Strehl and Ghosh (2002) used mutual information as a measure of the shared information between a pair of clusters. Based on a mutual information objective function, the authors evaluated the quality of different clusterings and the algorithm was able to select the best solution. In the field of clusters comparison, Vinh et al. (2010) proposed four different mutual information measures adjusted for chance and five distance measures that involved manipulation of entropy and mutual information measures as measures of comparison between clusters. According to the authors, the adjustment for chance is needed for the expected value of a similarity measure between pairs of independent clusters sampled independently at random to be constant. The distance measures were proposed as metric properties for the cluster comparisons. However, the authors did not discuss what is the specific motivation of each one of the distance measures.

Entropy, through the principle of maximum entropy and the principle of minimum cross-entropy, has been applied to decision trees and optimization models (X. Chen et al., 2012). In decision tree, fuzzy entropy was used to minimize the loss of information in a decision tree (X. Chen et al., 2012). Chandrasekaran and Shah (2017) studied relative entropy programs, which are optimization problems where the objective and the constraints are specified in terms of linear and relative entropy inequalities.

In multi-attribute decision making (MADM), Shannon's entropy has been used to find the weight of different criteria (Lotfi & Fallahnejad, 2010). In the MADM context, the greater the entropy value of an attribute, the smaller the attribute's weight, which

implies that the attribute has less discriminate power in the decision-making process. Asl, Khalilzadeh, Youshanlouei, and Mood (2012) used Shannon's entropy to rank different criteria, such as cost, software quality, vendor, and software capability, for selection of enterprise resource planning (ERP) systems within organizations.

Other applications of entropy through MaxEnt and MinxEnt include: buffer allocation, queueing models of telecommunication systems, scheduling, DNA sequence alignment, reinforcement learning, network reliability, astronomy, and signal processing (De Boer et al., 2005; Phillips et al., 2006; Shore & Johnson, 1981).

4.3. Material and methods

In this work, it is proposed to use entropy measures as a method for input parameter selection and experiment planning in simulation models. The ultimate goal is to answer the following research questions: (i) can entropy measures support the identification of the group of seeds that leads to the largest uncertainty, if any?; (ii) can entropy measures support the identification of the number of replications that leads to the largest uncertainty?; and, (iii) can entropy measures support the selection of the most important parameters?.

A 5-step procedure is followed to investigate the research questions of this work, as shown in Figure 75. Each step is described in detailed below.

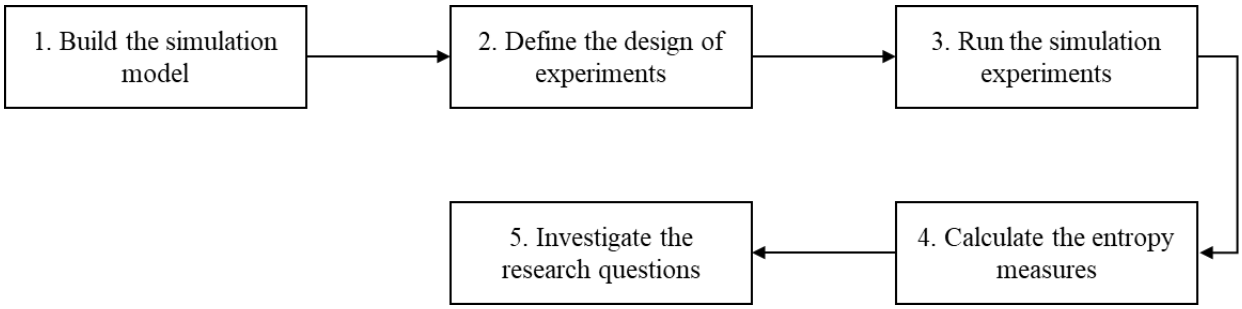

Figure 75. 5-step procedure to investigate the research questions.

4.3.1. Build the simulation model

The first step is to build the simulation model. Here, a single queue model was chosen for its simplicity to be used as an illustrative example and for having closed-form theoretical solutions for some distributions of the input parameters, as well as the output parameters of interest. The model consists of a single source of arrivals, a single queue, and s servers providing the same service. After being served, customers leave the system. Balking and reneging were not considered in the model. Two input parameters were considered in the model, namely inter-arrival time ($1/\lambda$) and service time ($1/\mu$), and two output responses were considered, namely average number of customers in the system (L or NIS) and average time spent in the system (W or TIS). The simulation generated input processes is also referred as $\vec{X} = [X_1$: arrival process, X_2: service process] and the output process as $\vec{Y} = [\hat{Y}_1$: average number of customers in the system, $\hat{Y}_2$: average time in the system]. The notation used here follows the $A/S/s$ Kendall's notation, where: A represents the arrival process, S the service time, and s the number of servers. M is used for memoryless distributions.

4.3.2. Define the design of experiments

The challenge of this work is to validate the appropriateness of entropy measures as a method of uncertainty quantification and, consequently, for identifying the experiment settings that lead to the largest uncertainty. To this purpose, it is important to use methods already exist and recognized in this field. Simulation uncertainty will be considered in two different ways: (i) $u_1 =$ standard deviation of the parameters of the experiments; and, (ii) $u_2 =$ number of parameter errors of the experiments, that is, number of simulations experiments where the simulated parameters were not within the theoretical confidence interval.

To be able to investigate whether different seeds, different number of replications, and the different input processes $\vec{X}$, full-factorial design was applied. Full-factorial was chosen to be able to identify whether the factors, as well as their interactions, would explain the variation in the simulation model. It is important to mention that seeds and number of replications were investigated separately, as they were the computer experiments replicates of each other. Three different seeds and three different values for each input process were investigated. The different values of the input processes were varied in a way that guaranteed that the highest traffic intensity of the queue model was around 90% to avoid instability of the model. Table 56 shows the different values of seeds and input processes considered in the experiments. Table 57 shows the full-factorial design of experiments.

Table 56. Parameter values for simulation experiments.

Parameter	Values

Interarrival time	$AT_1 = 20, AT_2 = 36, AT_3 = 180$
Service time	$ST_1 = 2, ST_2 = 10, ST_3 = 18.000018$
Interarrival time seed	$ATs_1 = 4, ATs_2 = 8, ATs_3 = 20$
Service time seed	$STs_1 = 3, STs_2 = 12, STs_3 = 15$
Number of replications	$r_1 = 10, r_2 = 20, r_3 = 50, r_4 = 100, r_5 = 200, r_6 = 400, r_7 = 600, r_8 = 800, r_9 = 1000, r_{10} = 1500$

Table 57. Full-factorial design of experiments. Each design was repeated for every replication $\boldsymbol{r_1}$ to $\boldsymbol{r_9}$.

Interarrival time	Service time	Interarrival time seed	Service time seed
AT_1	ST_1	ATs_1	STs_1
AT_1	ST_1	ATs_1	STs_2
AT_1	ST_1	ATs_1	STs_3
AT_1	ST_1	ATs_2	STs_1
AT_1	ST_1	ATs_2	STs_2
AT_1	ST_1	ATs_2	STs_3
AT_1	ST_1	ATs_3	STs_1
AT_1	ST_1	ATs_3	STs_2
AT_1	ST_1	ATs_3	STs_3
AT_1	ST_2	ATs_1	STs_1
AT_1	ST_2	ATs_1	STs_2
AT_1	ST_2	ATs_1	STs_3
AT_1	ST_2	ATs_2	STs_1
AT_1	ST_2	ATs_2	STs_2
AT_1	ST_2	ATs_2	STs_3
AT_1	ST_2	ATs_3	STs_1
AT_1	ST_2	ATs_3	STs_2
AT_1	ST_2	ATs_3	STs_3
AT_1	ST_3	ATs_1	STs_1
AT_1	ST_3	ATs_1	STs_2
AT_1	ST_3	ATs_1	STs_3
AT_1	ST_3	ATs_2	STs_1
AT_1	ST_3	ATs_2	STs_2
AT_1	ST_3	ATs_2	STs_3
AT_1	ST_3	ATs_3	STs_1
AT_1	ST_3	ATs_3	STs_2
AT_1	ST_3	ATs_3	STs_3
AT_2	ST_1	ATs_1	STs_1

AT_2	ST_1	ATs_1	STs_2
AT_2	ST_1	ATs_1	STs_3
AT_2	ST_1	ATs_2	STs_1
AT_2	ST_1	ATs_2	STs_2
AT_2	ST_1	ATs_2	STs_3
AT_2	ST_1	ATs_3	STs_1
AT_2	ST_1	ATs_3	STs_2
AT_2	ST_1	ATs_3	STs_3
AT_2	ST_2	ATs_1	STs_1
AT_2	ST_2	ATs_1	STs_2
AT_2	ST_2	ATs_1	STs_3
AT_2	ST_2	ATs_2	STs_1
AT_2	ST_2	ATs_2	STs_2
AT_2	ST_2	ATs_2	STs_3
AT_2	ST_2	ATs_3	STs_1
AT_2	ST_2	ATs_3	STs_2
AT_2	ST_2	ATs_3	STs_3
AT_2	ST_3	ATs_1	STs_1
AT_2	ST_3	ATs_1	STs_2
AT_2	ST_3	ATs_1	STs_3
AT_2	ST_3	ATs_2	STs_1
AT_2	ST_3	ATs_2	STs_2
AT_2	ST_3	ATs_2	STs_3
AT_2	ST_3	ATs_3	STs_1
AT_2	ST_3	ATs_3	STs_2
AT_2	ST_3	ATs_3	STs_3
AT_3	ST_1	ATs_1	STs_1
AT_3	ST_1	ATs_1	STs_2
AT_3	ST_1	ATs_1	STs_3
AT_3	ST_1	ATs_2	STs_1
AT_3	ST_1	ATs_2	STs_2
AT_3	ST_1	ATs_2	STs_3
AT_3	ST_1	ATs_3	STs_1
AT_3	ST_1	ATs_3	STs_2
AT_3	ST_1	ATs_3	STs_3
AT_3	ST_2	ATs_1	STs_1
AT_3	ST_2	ATs_1	STs_2
AT_3	ST_2	ATs_1	STs_3
AT_3	ST_2	ATs_2	STs_1
AT_3	ST_2	ATs_2	STs_2

AT_3	ST_2	ATs_2	STs_3
AT_3	ST_2	ATs_3	STs_1
AT_3	ST_2	ATs_3	STs_2
AT_3	ST_2	ATs_3	STs_3
AT_3	ST_3	ATs_1	STs_1
AT_3	ST_3	ATs_1	STs_2
AT_3	ST_3	ATs_1	STs_3
AT_3	ST_3	ATs_2	STs_1
AT_3	ST_3	ATs_2	STs_2
AT_3	ST_3	ATs_2	STs_3
AT_3	ST_3	ATs_3	STs_1
AT_3	ST_3	ATs_3	STs_2
AT_3	ST_3	ATs_3	STs_3

4.3.3. Run the simulation experiments

The third step was to run the simulation experiments. The experiments were run in Simio® University Enterprise Edition v 12.205. The experiment was run for 1825 days, including a warm-up period of 365 days. Two different queue systems were considered: (i) $M/M/1$, and (ii) $M/M/inf$. As described in step 2, the experiments included different system configurations by varying the parameter values and, consequently, the traffic intensities, the seeds for generating random numbers, and the number of replications. This led to a total of 1620 experiments.

4.3.4. Calculate the entropy measures

The next step is to apply entropy measures to quantify the uncertainty of the simulation outputs ($H(Y)$) and the uncertainty of the simulation generated inputs ($H(X)$). The entropy measures were calculated using two different approaches: (i) histogram discrete empirical estimate, and (ii) histogram probability density estimate. The second is

a theoretically correct approach but its application is challenging due to the difficulties that arise when computing entropy measures for continuous variables. These difficulties were discussed in section 2 and section 3. The first approach has practical benefits in terms of calculations and it is frequently used by practitioners. The first and second approaches were used in section 2 to investigate the potential of entropy measures as a measure of uncertainty quantification in simulation models. As pointed out in section 2, despite not being the theoretically correct approach, the entropy measures calculated using the histogram discrete empirical estimate presented better results in capturing the deterministic behavior of the system when compared to the histogram probability density estimate. Therefore, using the approach here is justified.

In order to handle some of the challenges encountered while applying entropy measures for continuous variables, the differential entropy measures in this work were calculated using Equation 79, which was proposed by Jaynes (1962), and approximated through Equation 80, given by Xiong et al. (2017) and Steuer et al. (2002).

$$H(X) = -\int p(x) \log\left[\frac{p(x)}{m(x)}\right] dx \qquad \text{Equation 79}$$

$$\widehat{H}(X) = -\frac{1}{n}\sum_{i=1}^{n} \log \hat{f}(X_i) \qquad \text{Equation 80}$$

In Jaynes' equation it is important to define the invariant measure $m(x)$. Here, $m(x) = \hat{f}(x)\big(1 + \hat{f}(x)\big)$, based on what was proposed in section 3.

4.3.5. Investigate the research questions

The fifth step is to investigate the three research questions: (i) can entropy measures support the identification of the group of seeds that leads to the largest uncertainty, if any?; (ii) can entropy measures support the identification of the number of replications that leads to the largest uncertainty?; and, (iii) can entropy measures support the selection of the most important parameters?.

The queue model used in this work consists of two input parameters (i.e. X_1 – interarrival time and X_2 – service time) and two output responses (i.e. $\hat{Y}_1$ - average number of customers in the system and $\hat{Y}_2$ - average time in the system). Consequently, to answer question (i) it is important to investigate for each of the inputs and outputs whether there are statistical significantly differences in the entropy measures based on the seeds used on the experiments and whether these results are consistent with another measure of uncertainty. Standard error of the mean (SEM) is used as the uncertainty measure for comparison. SEM quantifies uncertainty by measuring how far the sample mean of the data is likely to be from the true population mean (Altman & Bland, 2005; Barde & Barde, 2012). SEM is widely recognized in the academic community as a measure of uncertainty or precision of the mean and is used as a means of calculating confidence intervals. To identify the statistical differences, regression analysis and Tukey-Kramer multiple comparison test using JMP® Pro 15 is used.

Question (ii) is answered in a similar way to question (i). Regression analysis and Tukey-Kramer multiple comparison test is used to investigate for each of the inputs and outputs whether there are statistical significantly differences in the entropy measures

based on the number of replications used on the experiments and whether these results are consistent with the results obtained from the standard error of the mean.

To answer question (iii), it is important to investigate what are the most important parameters for the simulation model based on the entropy measures and whether these results are consistent with another known method. As a known method in the scientific community, ANOVA is used to identify what are the significant parameters for each of the simulation responses. Mutual information measures is used to identify what are the most important parameters for each simulation response. More specifically, in this work, it is proposed to use an adaptation of the mutual information index for this purpose. As discussed in section 2.1, the mutual information index is given by $MII_{X_1,Y} = \frac{I(X_1;Y)}{I(Y,Y)} = \frac{I(X_1;Y)}{H(Y)}$. The measure can be calculated and compared for different inputs to determine which input provide useful information about the output. Contigency analysis is used to test the significance of the results.

A limitation of $MII_{X_j;Y_i}$ is that in a model with only two input parameters X_1 and X_2, $MII_{X_1,Y_i} + MII_{X_1,Y_i}$ does not add to 1. To overcome this limitation, it is proposed an adaptation of the mutual information index. For the $M/M/s$ models, it is known that the inputs X_1 and X_2 are independent. Hence: $I(X_1;X_2) = H(X_1) - H(X_1|X_2) = H(X_1) - H(X_1) = 0$. Because $I(X_1;X_2) = 0$, we can calculate the total mutual information of Y_i ($TI(Y_\text{i})$) using Equation 81:

$$TI(Y_i) = I(Y_i;X_1) + I(Y_i;X_2) \qquad \text{Equation 81}$$

Now, the impact of input X_j on output Y_i ($IIO_{X_j;Y_i}$) can be calculated using Equation 82:

$$IIO_{X_j;Y_i} = \frac{I(Y_i; X_j)}{TI(Y_i)} \qquad \text{Equation 82}$$

$IIO_{X_j;Y_i}$ leads to same conclusions as $MII_{X_j;Y_i}$ with respect to the input that provides more useful information about the output. However, $IIO_{X_j;Y_i}$ has the advantage that will add to 1. On the other hand, $IIO_{X_j;Y_i}$ can only be applied when the inputs are independent.

After the $IIO_{X_j;Y_i}$ was calculated for each output of the simulation, it was identified which input provided more useful information on each response. A value of 1 was used as an indicator for this and 0 otherwise. For cases where there was a tie, both inputs received 1. Contingency analysis was performed in JMP® Pro15 to test whether one of the inputs provided statistically significantly more useful information than the other. The results obtained from the contingency analysis was compared to the ANOVA results. Contingency analysis was also used to test whether the useful information provided by each input to the output varied based on different factors, such as: seed used for the inputs, number of replications, queue model, and traffic intensity.

As a final note on the methods, it is important to mention that validating the results of the entropy measures is not a simple task, as other methods of uncertainty quantification in simulation models have different goals and use different approaches that are difficult to compare, such as the work of Xie et al. (2014a), which estimated the uncertainty of the input models resulting from using limited real data. This uncertainty was used to update

the Bayesian credible interval. Therefore, in an attempt to validate the results of this work and minimize its limitations, the results obtained from the entropy measures were compared with results from well-known methods in the literature that have a goal as similar as possible.

4.4. Results and discussion

4.4.1. Regression model for the SEM of each input and output

To investigate whether the different random seeds used to generate the inputs of the simulation model resulted in different uncertainty in the simulated inputs and outputs, a second-order regression model for the SEM of each input and each output was built. The following were considered as possible factors for the regression models of the SEM: (i) seed used in the interarrival time input X_1; (ii) seed used in the service time input X_2; (iii) interarrival time mean value (X_1); (iv) service time mean value (X_2); (v) queue model (i.e., $\mathrm{M/M/1}$ or $\mathrm{M/M/inf}$); (vi) traffic intensity; and, (vii) number of replications.

For each of the input and output, some of the factors were collinear or aliased (e.g., traffic intensity) and could not be estimated independently and others were not statistically significant (e.g., queue model). After performing the analysis, a second-order linear regression model with different independent factors was fitted to each dependent factor, i.e., the SEM of each simulation input and output, using least squares approach. The summary of fit of each model, as well as the independent factors included in the model are shown in Table 58. Figure 76 and Figure 77 show the actual by predicted plot of the models.

Table 58. Summary of fit of SEM regression models and summary of effects included in the models for α-level $= \mathbf{0.05}$.

Model	**SEM X_1**	**SEM X_2**	**SEM $\hat{Y}_1$**	**SEM $\hat{Y}_2$**
Model p-value	<0.0001	<0.0001	<0.0001	<0.0001
R^2	0.9978	0.9941	0.9907	0.9837
$R^2 adj$	0.9978	0.9941	0.9906	0.9836
Factor	**p-value**	**p-value**	**p-value**	**p-value**
Seed interarrival time	0.0000	N/A	0.0000	0.0000
Interarrival time mean value	0.0000	0.0000	N/A	N/A
Seed service time	N/A	0.0025	0.0182	0.0275
Service time mean value	N/A	0.0000	N/A	N/A
Number of replications	0.0000	0.0000	0.0000	0.0000
Traffic intensity	N/A	N/A	0.0000	0.0000
Seed interarrival time * Interarrival time mean value	0.0000	N/A	N/A	N/A
Seed interarrival time * Number of replications	0.0000	N/A	N/A	0.0452
Interarrival time mean value * Number of replications	0.0000	0.0000	N/A	N/A
Interarrival time mean value * Seed service time	N/A	0.0000	N/A	N/A
Interarrival time mean value * Service time mean value	N/A	0.0000	N/A	N/A
Seed service time * Service time mean value	N/A	N/A	N/A	N/A
Seed service time * Number of replications	N/A	0.0000	N/A	N/A
Service time mean value * Number of replications	N/A	0.0000	N/A	N/A
Seed interarrival time * Seed service time	N/A	N/A	0.0481	0.0494
Number of replications * Traffic intensity	N/A	N/A	0.0000	0.0000
Seed interarrival time * Traffic intensity	N/A	N/A	0.0000	0.0000
Seed service time * Traffic intensity	N/A	N/A	0.0000	0.0000

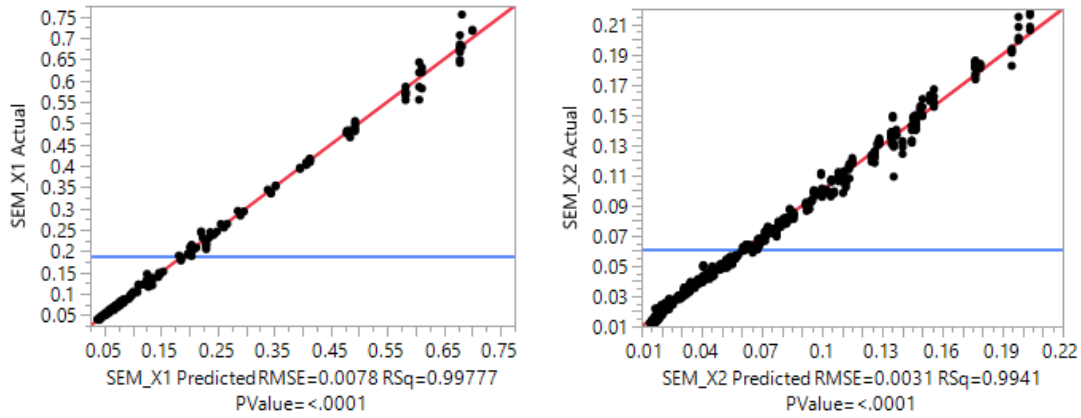

Figure 76. Actual by predicted plot of the standard error of the mean of the inputs. Left-side: $\boldsymbol{X_1}$ - interarrival time. Right-side: $\boldsymbol{X_2}$ - service time.

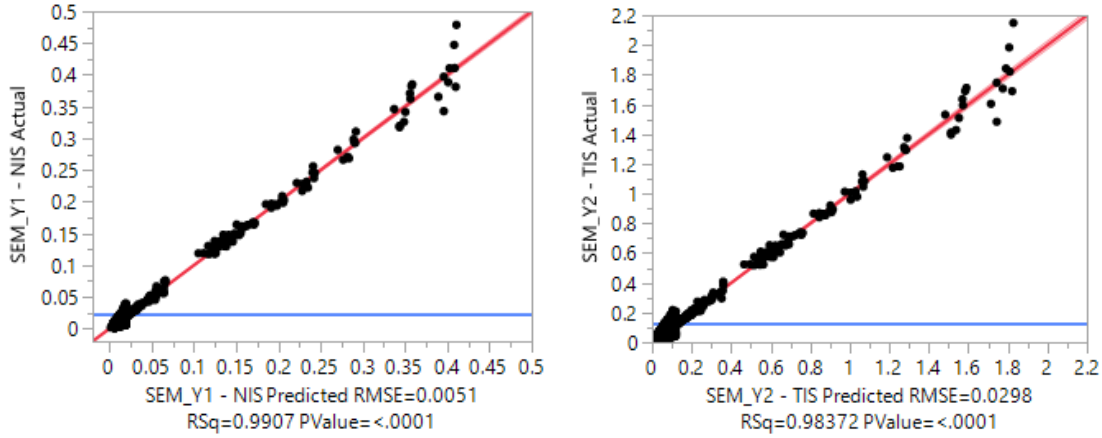

Figure 77. Actual by predicted plot of the standard error of the mean of the outputs. Left-side: $\boldsymbol{\hat{Y}_1}$ – average number in system. Right-side: $\boldsymbol{\hat{Y}_2}$ – average time in system.

4.4.2. Regression model for the entropy measures of each input and output

Next, a second-order regression model for the entropy of each input and each output was built. This was done for the entropy calculated by each of the two different approaches: (a) histogram discrete empirical estimate, and (b) histogram probability density estimate. Therefore, a total of 8 regression models were fitted: 2 approaches x (2

simulation inputs + 2 simulation outputs) = 8. For the entropy model, an extra independent factor was also considered: number of bins. Consequently, the following different factors were considered as possible factors for the regression models: (i) seed used in the interarrival time input X_1; (ii) seed used in the service time input X_2; (iii) interarrival time mean value (X_1); (iv) service time mean value (X_2); (v) queue model (i.e. $M/M/1$ or $M/M/inf$); (vi) traffic intensity; (vii) number of replications; and, (viii) number of bins.

In all the regression models, some of the factors were also aliased and could not be estimated independently or they were not statistically significant and, consequently, were not included in the model. After performing the analysis, a linear regression model with different independent factors was fitted to each dependent factor, i.e., the entropy of each simulation input and output, using least squares approach. The summary of fit of each model, as well as the independent factors included in the model is shown next.

(a) Entropy measures calculated using the histogram discrete empirical estimate:

The summary of fit of each model and the independent factors included in the model for both the entropy measures and the normalized entropy measures calculated using the histogram discrete empirical estimate are shown in Table 59. Figure 78 and Figure 79 show the actual by predicted plot of the models of the entropy measures as an example.

Table 59. Summary of fit of the regression models for the entropy measures calculated using the histogram discrete empirical estimate and summary of effects included in the models for α-level $= \mathbf{0.05}$.

Model	$\boldsymbol{H(X_1)}$	$\boldsymbol{H(X_2)}$	$\boldsymbol{H(Y_1)}$	$\boldsymbol{H(Y_2)}$	**Norm** $\boldsymbol{H(X_1)}$	**Norm** $\boldsymbol{H(X_2)}$	**Norm** $\boldsymbol{H(Y_1)}$	**Norm** $\boldsymbol{H(Y_2)}$

Model p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R^2	0.9993	0.9987	0.9982	0.9983	0.9848	0.9535	0.9482	0.9452
$R^2 adj$	0.9992	0.9986	0.9981	0.9983	0.9847	0.9530	0.9474	0.9443
Factor	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**
Number of clusters	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed interarrival time	0.0000	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000
Interarrival time mean value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed service time	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000
Service time mean value	N/A	N/A	0.0021	0.0000	N/A	N/A	0.0000	0.0000
Number of replications	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Number of bins * Seed interarrival time	0.0000	N/A	0.0000	N/A	0.0000	N/A	0.0000	N/A
Number of bins * Interarrival time mean value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Number of bins * Seed service time	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000
Number of bins * Service time mean value	N/A	N/A	N/A	N/A	N/A	N/A	0.0000	0.0000
Number of bins * Number of replications	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed interarrival time * Interarrival time mean value	0.0000	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0052
Seed interarrival time * Seed service time	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0000	0.0001
Seed interarrival time * Service time mean value	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0018	0.0003

Seed interarrival time * Number of replications	0.0000	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0087
Seed interarrival time * Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Interarrival time mean value * Seed service time	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0000	N/A
Interarrival time mean value * Seed service time	N/A	0.0000	N/A	N/A	N/A	0.0000	N/A	0.0000
Interarrival time mean value * Service time mean value	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0000	0.0000
Interarrival time mean value * Number of replications	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed service time * Service time mean value	N/A	N/A	0.0004	0.0003	N/A	N/A	N/A	N/A
Seed service time * Number of replications	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000
Service time mean value * Number of replications	N/A	N/A	N/A	0.0000	N/A	N/A	N/A	0.0000
Seed service time * Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Number of replications * Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

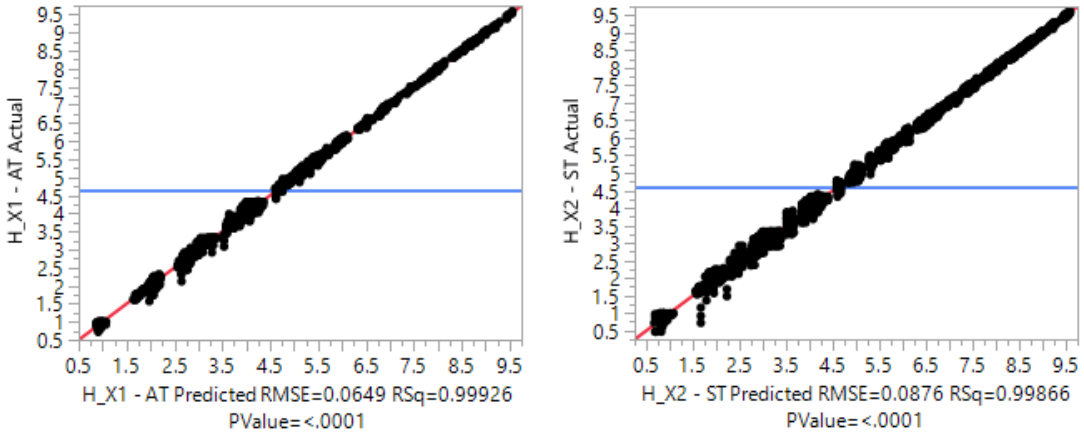

Figure 78. Actual by predicted plot of the entropy measures of the inputs calculated using the histogram discrete empirical estimate. Left: $\boldsymbol{H(X_1)}$ – entropy of interarrival time. Right: $\boldsymbol{H(X_2)}$ – entropy of service time.

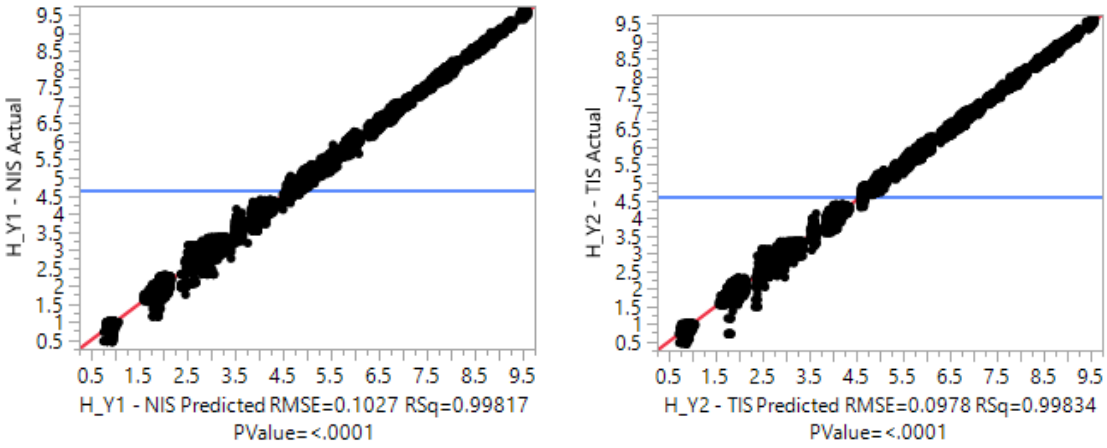

Figure 79. Actual by predicted plot of the entropy measures of the outputs calculated using the histogram discrete empirical estimate. Left-side: $\boldsymbol{H(Y_1)}$ – entropy of number in system. Right-side: $\boldsymbol{H(Y_2)}$ – entropy of time in system.

(b) Entropy measures calculated using the histogram probability density estimate:

The summary of fit of each model and the independent factors included in the model for both the entropy measures and the normalized entropy measures calculated using the histogram probability density estimate are shown in Table 70.

Table 60. Summary of fit of the regression models for the entropy measures calculated using the histogram probability density estimate and summary of effects included in the models for α-level $= \mathbf{0.05}$.

Model	$\boldsymbol{H(X_1)}$	$\boldsymbol{H(X_2)}$	$\boldsymbol{H(Y_1)}$	$\boldsymbol{H(Y_2)}$	**Norm** $\boldsymbol{H(X_1)}$	**Norm** $\boldsymbol{H(X_2)}$	**Norm** $\boldsymbol{H(Y_1)}$	**Norm** $\boldsymbol{H(Y_2)}$
Model p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R^2	0.9974	0.9952	0.9933	0.9939	0.9998	0.9992	0.9991	0.9991
$R^2 adj$	0.9974	0.9951	0.9932	0.9938	0.9998	0.9992	0.9991	0.9991
Factor	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**	**p-value**
Number of bins	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed interarrival time	0.0000	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000
Interarrival time mean value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed service time	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000
Service time mean value	N/A	N/A	0.0006	0.0000	N/A	N/A	0.0000	0.0000
Number of replications	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Number of bins * Seed interarrival time	0.0000	N/A	0.0000	N/A	0.0000	N/A	0.0000	N/A
Number of bins * Interarrival time mean value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Number of bins * Seed service time	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000
Number of bins * Service time mean value	N/A	N/A	N/A	N/A	N/A	N/A	0.0000	0.0000

Number of bins * Number of replications	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed interarrival time * Interarrival time mean value	0.0000	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0066
Seed interarrival time * Seed service time	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0000	0.0010
Seed interarrival time * Service time mean value	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0028	0.0007
Seed interarrival time * Number of replications	0.0000	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0384
Seed interarrival time * Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Interarrival time mean value * Seed service time	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0000	N/A
Interarrival time mean value * Seed service time	N/A	0.0000	N/A	N/A	N/A	0.0000	N/A	0.0000
Interarrival time mean value * Service time mean value	N/A	N/A	0.0000	0.0000	N/A	N/A	0.0000	0.0000
Interarrival time mean value * Number of replications	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Seed service time * Service time mean value	N/A	N/A	0.0003	0.0007	N/A	N/A	N/A	N/A
Seed service time * Number of replications	N/A	0.0000	0.0000	0.0000	N/A	0.0000	0.0000	0.0000
Service time mean value * Number of replications	N/A	N/A	N/A	0.0000	N/A	N/A	N/A	0.0000

Seed service time * Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Number of replications * Traffic intensity	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

4.4.3. Can entropy measures support the identification of the group of seeds that leads to the largest uncertainty, if any?

To answer whether entropy can support the identification of the group of seeds that leads to the largest uncertainty, first, it is important to understand whether different seeds lead to statistically significantly different uncertainty in the model according to a baseline measure of uncertainty. In this work, the baseline measure of uncertainty is the standard error of the mean. For this, the linear regression model of the SEM of each input and output discussed in section 4.4.1 was used together with Tukey-Kramer multiple comparison test. Next, it is important to investigate whether different seeds lead to statistically significantly different values of entropy. For this, the linear regression model of the entropy measures of the inputs and outputs, from section 4.4.2, together with Tukey-Kramer multiple comparison test were used. The final step is to compare the results to understand whether the results from the entropy measures are consistent with the results from the standard error of the mean or not. Table 71 shows the results of the Tukey-Kramer multiple comparison test for the linear regression model of the SEM and the entropy measures. The results are shown as a connecting letter report, where any levels that share a letter do not have a statistically significant difference.

First, the results in Table 68, Table 69, and Table 70 show that seed is a statistically significantly factor for both the SEM and the entropy measures of all the inputs and outputs. Next, Tukey-Kramer multiple comparison test was used to understand the difference among the seeds. The following important observations can be made from Table 71:

1. With respect to the statistical differences between the seeds, in general, the results from the entropy measures are similar to the results of the SEM. When the results are different, the entropy measures appear to be more sensitive to differences and are able to capture statistically significant differences more frequently than the SEM.
2. Although the discrete empirical distribution estimate is not the theoretically correct approach, as previously discussed, the estimate presented better results in capturing the deterministic behavior of the simulated queue system. However, with a few exceptions, with respect to identifying statistical differences between the seeds, the discrete empirical distribution estimate showed results very similar to the probability density estimate method. This is an indicator that the discrete empirical distribution estimate could be, in fact, a potential estimate for measuring uncertainty in simulation models.
3. The results from the normalized and non-normalized version of the entropy measures are also similar, except for a few cases. This indicates that when

the goal is to identify the group of seeds that lead to the largest uncertainty, normalization may not be necessary.

4. According to the entropy measures, NIS appears to be more sensitive to the seeds than the TIS, as there were more statistically significant differences for the former than the latter.

5. To identify the seed that leads to the largest uncertainty, the least square mean of each measure was ordered from the smallest to the largest, as shown in Table 71. At first, the entropy measures do not appear to show good results. The entropy measures results are not consistent with the results of the SEM. To investigate the results in more detail, it was decided to compare the results of the entropy measures and the SEM with the confidence interval of the inputs and outputs from the simulated queue model. For every simulated scenario, if the theoretical values of the inputs and outputs were within their respective confidence interval, it was not considered an error. If the theoretical values were not within the confidence interval, it was considered as an error. Figure 80, Figure 81, and Figure 82 show the graphs of the total number of errors per seed for each input and output. When one compares the results of Table 71 with the results in Figure 80, Figure 81, and Figure 82, it is possible to see that neither methods, that is, nor the entropy measures, nor the SEM, are able to identify the seed that leads to the largest uncertainty consistently with the seed that leads to the highest number of errors. It is important to highlight

that as shown in Table 69 and Table 70, there are interactions among the seed and other parameters, such as number of replications, number of bins, interarrival time, etc. Despite these interactions, one would expect that overall, the methods would be consistent. However, for the entropy measures it is known that the number of bins has a great impact on the entropy measures values and it can be seen as an overfitting or underfitting parameter. When specific parameter values were selected, for instance, number of replications = 50, interarrival time = 180 min, and number of bins = 1000, the entropy measure was able to consistently detect the seed that led to the largest number of errors, as shown in Figure 83. This highlights the importance of adequately selecting the proper number of bins for the entropy measures. As a future work in this area, one could explore optimization models to search for the optimum number of bins, where the objective function would be to have the greatest consistency between the seed identified by the entropy measures and the seed identified by the largest number of errors. Another suggestion in this area is to use the conditional entropy, instead of the entropy as the measure of the uncertainty. The entropy is the average total uncertainty of the input or output, but it also contains the amount of information contained in the input that helps predict the output (or vice-versa). Therefore, a correlation is expected with measures of error, but not necessarily a match. The conditional entropy could be a better measure for these cases where one

wants to eliminate the amount of information provided by other variables, as shown in Figure 84.

Table 61. Tukey-Kramer multiple comparison test results for seed parameter for α-level $= 0.05$.

Input or Output	Seed	Method	Measure	Connecting letter report			Least Square Mean order		
				Seed1	Seed2	Seed3	Seed1	Seed2	Seed3
AT	Seed AT		SEM	A	A	B	1	2	3
		Discrete	H	A	C	B	1	3	2
			Norm H	A	C	B	1	3	2
		PDF	H	A	C	B	1	3	2
			Norm H	A	C	B	1	3	2
ST	Seed ST		SEM	A	A,B	B	3	2	1
		Discrete	H	A	C	B	3	1	2
			Norm H	A	C	B	3	1	2
		PDF	H	A	C	B	3	1	2
			Norm H	A	C	B	3	1	2
NIS	Seed AT		SEM	A	A	B	2	3	1
		Discrete	H	A	C	B	1	3	2
			Norm H	A	C	B	1	3	2
		PDF	H	A	C	B	1	3	2
			Norm H	A	C	B	1	3	2
	Seed ST		SEM	A	B	A,B	3	1	2
		Discrete	H	A	B	C	1	3	2
			Norm H	A	B	C	1	3	2
		PDF	H	A	B	A	1	3	2
			Norm H	A	B	C	1	3	2
TIS	Seed AT		SEM	A	A	B	2	3	1
		Discrete	H	A	B	B	1	2	3
			Norm H	A	B	B	1	2	3
		PDF	H	A	C	B	1	2	3
			Norm H	A	B	B	1	2	3
	Seed ST		SEM	A	B	A,B	3	1	2
		Discrete	H	A	B	B	3	1	2
			Norm H	A	B	B	3	1	2
		PDF	H	A	B	B	3	1	2
			Norm H	A	B	B	3	1	2

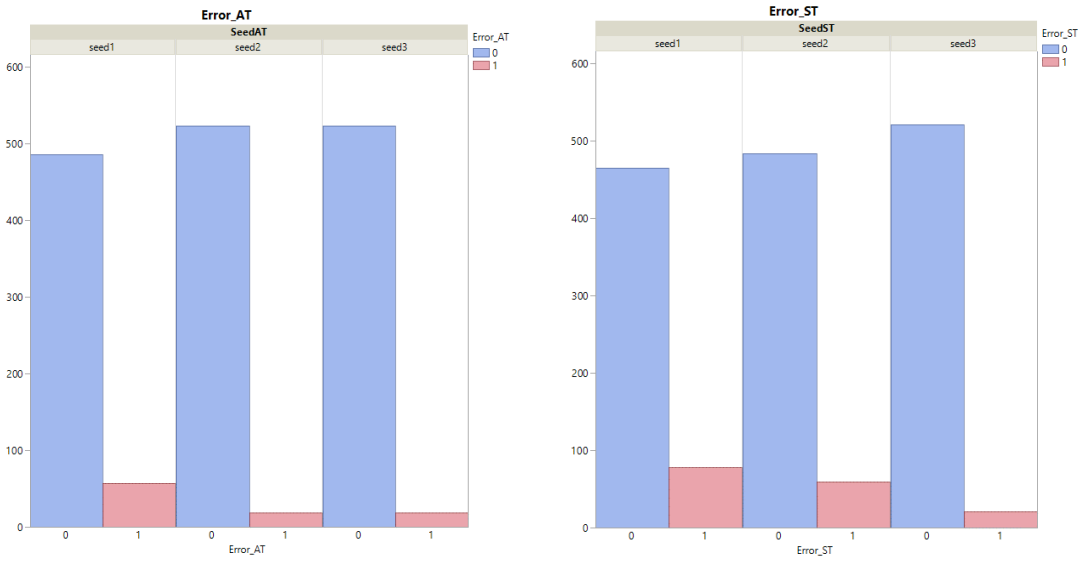

Figure 80. Total number of errors of the inputs interarrival time and service time per seed. Left-side: Total number of errors in interarrival time per seed used. Right-side: Total number of errors in service time per seed used.

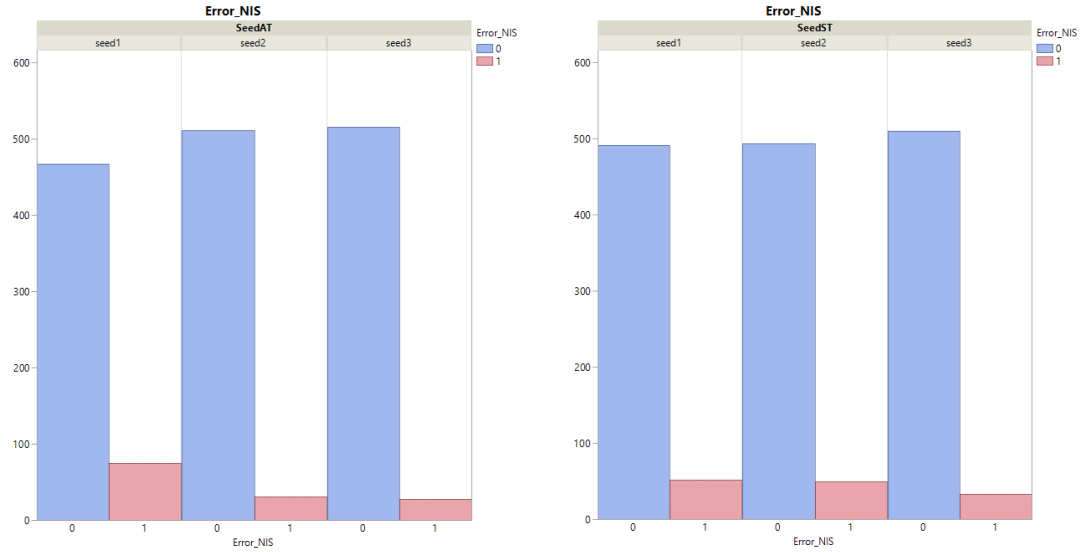

Figure 81. Total number of errors of the output number in system per seed. Left-side: Total number of errors in number in system per interarrival time seed used. Right-side: Total number of errors in number in system per service time seed used.

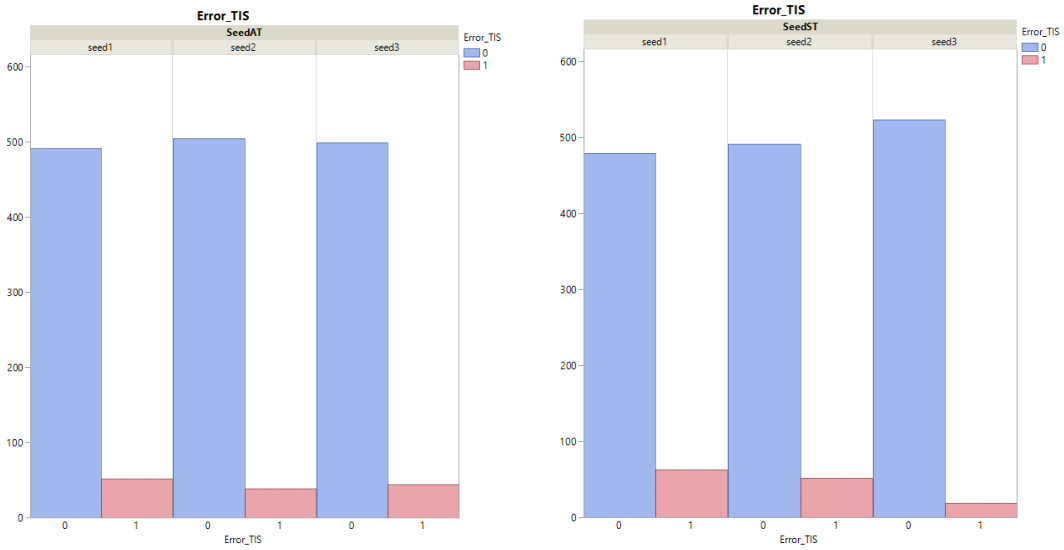

Figure 82. Total number of errors of the output time system per seed. Left-side: Total number of errors in time in system per interarrival time seed used. Right-side: Total number of errors in time in system per service time seed used.

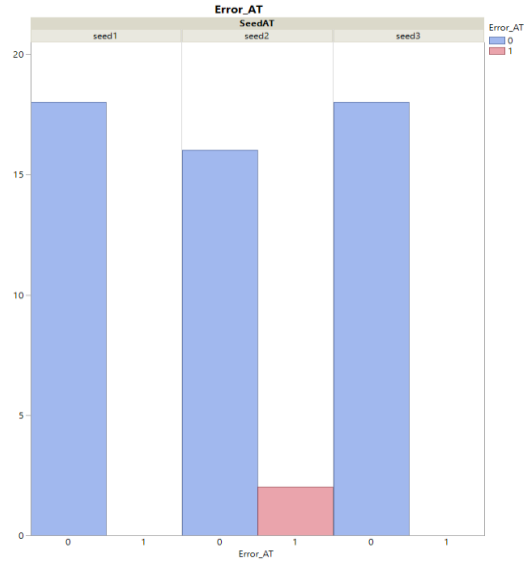

Figure 83. Total number of errors of the input interarrival time per seed for number of replications = 50 and interarrival time = 180 minutes.

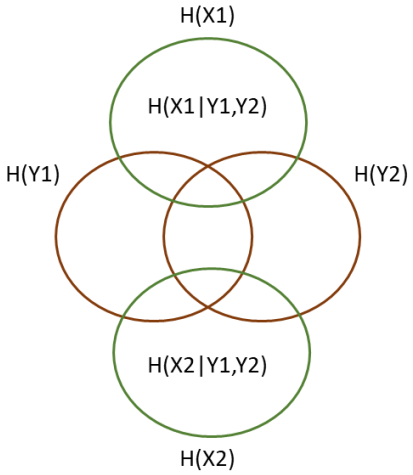

Figure 84. Conditional entropy $\boldsymbol{H(X_i|Y_i,Y_j)}$.

4.4.4. Can entropy measures support the identification of the number of replications that leads to the largest uncertainty?

To answer whether entropy can support the identification of the number of replications that leads to the largest uncertainty, first, it is important to understand whether different replications lead to statistically significantly different uncertainty in the model according to a baseline measure of uncertainty. For this, the linear regression model of the SEM of each input and output discussed in section 4.4.1 was used together with Tukey-Kramer multiple comparison test. Next, it is important to investigate whether different number of replications lead to statistically significantly different values of entropy. For this, the linear regression model of the entropy measures of the inputs and outputs, from section 4.4.3, together with Tukey-Kramer multiple comparison test were used. The final step is to compare the results to understand whether the results from the entropy measures are consistent with the results from the standard error of the mean or not. Table 72 shows

the results of the Tukey-Kramer multiple comparison test for the linear regression model of the SEM and the entropy measures. The results are shown as a connecting letter report, where any levels that share a letter do not have a statistically significant difference.

First, the results in Table 68, Table 69, and Table 70 show that number of replications is a statistically significant factor for both the SEM and the entropy measures of all the inputs and outputs. Next, Tukey-Kramer multiple comparison test was used to understand the difference among the number of replications. The following important observations can be made from Table 72:

1. With respect to the statistical differences between the replications, in general, the results from the entropy measures are similar to the results of the SEM. If one considers the results of the non-normalized version of the entropy measures, the entropy measures appear to be more sensitive to differences among the replications and be able to capture statistically significant differences more frequently than the SEM. When one considers the normalized version, the results of the entropy measures and SEM are very similar. The exception is for the entropy of interarrival time, where the entropy measure appeared to be less sensitive than the SEM. However, in this case, the normalized version appears to be a better measure because it places all the entropy measures within the same range between 0 and 1, which makes the comparison among themselves more adequate. Without the normalization, the entropy measures would have different range values depending on the number of bins used to calculate the entropy.

2. Considering the normalized measures, all three methods showed results very similar to each other. Especially, with respect to the number of replications equal to 800 and 1,000, all three methods were not able to identify any statistical significance between them, which indicates that running 1,000 replications may potentially not be worth it if one is constrained by computer resources and/or time. For entropy of interarrival time, the entropy measures were not able to identify statistically significant differences not even between 400 replications and 800 replications.
3. According to the entropy measures, the inputs appear to be more sensitive to the replications than the outputs, as there were more statistically significant differences for the first group than the latter.
4. As the results in Table 72 show, while the SEM decreases with the increase in the number of replications, the entropy increases. At first, one could think that this would indicate that the entropy measures results are not consistent with the results of the SEM. However, this is not the case. The SEM is a measure divided by the number of samples, in this case the number of replications. Therefore, it is expected that as the number of replications increase, SEM should decrease. Entropy, on the other hand, is the average total uncertainty or information content. Intuitively, as the number of replications increase, it is expected that the total uncertainty or information content will also increase. A more technical explanation is that entropy is a function that decreases as the probability of an event increases.

In general, with more replications, more events may be observed and the probability of an event may in fact decrease. Therefore, if one wants to choose the best number of replications, one cannot simply use the total entropy, but a suggestion would be to calculate the entropy per replication or to calculate the gain in entropy per replication, which can be calculated from a set number of replications to another number of replications, such as $H(X_{set2}) - H(X_{set1})/(nreps_{set2} - nreps_{set1})$.

Table 62. Tukey-Kramer multiple comparison test results for number of replications parameter for α-level = $\mathbf{0.05}$.

Input or Output	Method	Measure	Number of replications connecting letter report									
			10	**20**	**50**	**100**	**200**	**400**	**600**	**800**	**1000**	**1500**
AT	SEM		A	B	C	D	E	F	G	H	H	I
	Discrete	H	J	I	H	G	F	E	D	C	B	A
		Norm H	G	F	E	D	C	A,B	B	A	B	A,B
	PDF	H	I	H	G	F	E	D	C	B	B	A
		Norm H	G	F	E	D	C	A,B	B	A	B	A,B
ST	SEM		A	B	C	D	E	F	G	G,H	H,I	I
	Discrete	H	J	I	H	G	F	E	D	C	B	A
		Norm H	I	H	G	F	E	D	C	B	A,B	A
	PDF	H	J	I	H	G	F	E	D	C	B	A
		Norm H	I	H	G	F	E	D	C	B	A,B	A
NIS	SEM		A	B	C	D	E	F	F,G	F,G	F,G	G
	Discrete	H	J	I	H	G	F	E	D	C	B	A
		Norm H	H	G	F	E	D	C	B	A,B	A	A
	PDF	H	J	I	H	G	F	E	D	C	B	A
		Norm H	H	G	F	E	D	C	B	A,B	A,B	A
TIS	SEM		A	B	C	D	E	F	F,G	F,G	F,G	G
	Discrete	H	J	I	H	G	F	E	D	C	B	A
		Norm H	H	G	F	E	D	C	B	A	A	A
	PDF	H	J	I	H	G	F	E	D	C	B	A
		Norm H	H	G	F	E	D	C	B	A	A	A

4.4.5. Can entropy measures support the selection of the most important parameters?

To answer the question whether entropy measures can support the selection of the most important input parameters in a simulation model, first it is important to investigate what are the significant parameters for the simulation model according to a known method in the literature. In this work, ANOVA is used to identify the significant parameters for each simulation response. Next, it is important to investigate what are the most important parameters based on the mutual information measures. For this, the mutual information index, as described in section 4.3.5, is used. The results obtained from the mutual information index was tested using contingency analysis. Finally, the results from the mutual information index were compared to the results of ANOVA.

For the ANOVA, the following different factors were considered: (i) interarrival time mean value (X_1); (ii) service time mean value (X_2); (iii) seed interarrival time; (iv) seed service time; (v) traffic intensity; and, (vii) number of replications. The results of ANOVA are shown in Table 63. Figure 85 shows the actual by predicted plot of the models of the simulation responses. From the results in Table 63, one can see that both the interarrival time and the service time are considered statistically significant factors for the simulation responses NIS and TIS.

Table 63. ANOVA results for simulation responses for α-level $= \mathbf{0.05}$.

Model	**NIS**	**TIS**
Model p-value	<0.0001	<0.0001
R^2	0.9999	0.9999

R^2adj	0.9999	0.9999
Factor	**p-value**	**p-value**
Seed interarrival time	0.0003	0.0012
Interarrival time mean value	0.0000	0.0066
Seed service time	N/A	N/A
Service time mean value	0.0000	0.0000
Number of replications	0.0017	0.0001
Traffic intensity	0.0000	0.0000
Seed interarrival time * Interarrival time mean value	0.0000	0.0000
Seed interarrival time * Seed service time	N/A	N/A
Seed interarrival time * Service time mean value	0.0000	0.0000
Seed interarrival time * Number of replications	0.0004	0.0040
Seed interarrival time * Traffic intensity	N/A	N/A
Interarrival time mean value * Seed service time	N/A	N/A
Interarrival time mean value * Seed service time	N/A	N/A
Interarrival time mean value * Service time mean value	0.0000	N/A
Interarrival time mean value * Number of replications	0.0017	0.0011
Seed service time * Service time mean value	N/A	N/A
Seed service time * Number of replications	N/A	N/A
Service time mean value * Number of replications	0.0002	0.0001
Seed service time * Traffic intensity	N/A	N/A
Number of replications * Traffic intensity	N/A	N/A

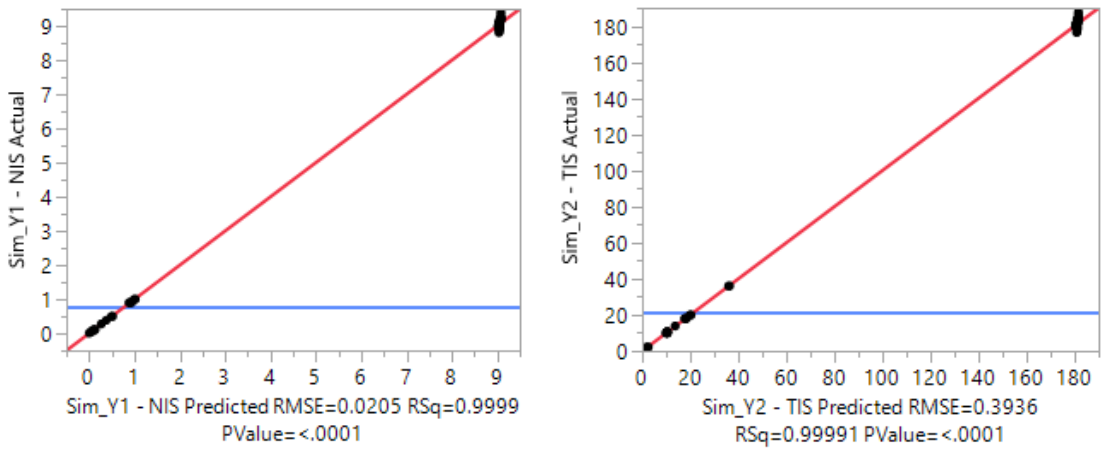

Figure 85. Actual by predicted plot of the simulation responses. Left-side: $\widehat{\boldsymbol{Y}}_{\mathbf{1}}$ - NIS. Right-side: $\widehat{\boldsymbol{Y}}_{\mathbf{2}}$ - TIS.

Contingency analysis was performed in JMP® Pro15 to test whether one of the inputs (e.g., interarrival time) provided statistically significantly more information than the other (e.g., service time). The results from the contingency analysis are summarized in Table 64 and shown in Figure 86 and Figure 87. Figure 86 and Figure 87 show the results for the non-normalized MI only. Contingency analysis was also used to test whether the useful information provided by each input to the output varied based on different factors, such as: seed used for the inputs, number of replications, queue model, and traffic intensity. These results are summarized in Table 65.

The following important observations can be made from Table 64 with respect to the importance of the inputs to the outputs:

1. The normalization does not appear to have an impact on the importance of the inputs to the outputs. The results were the same regardless of using normalization or not. Similar observation can be made from Table 65.
2. The results from the MI estimation method were consistent for the output NIS, but differed for the output TIS. Based on these results, it is difficulty to validate whether one method is better than the other or whether one method is correctly estimating the most significant input at all. According to Law (2007), time in system appears to be more impacted by service time than by arrival time. Assuming this is correct, the discrete empirical estimate is leading to better results than PDF, which matches the results from section 2.4. This also highlights the importance of choosing the responses of interest when modeling a system. The most significant

parameters in a model may be different depending on the response of interest. When a modeler is constrained by time and resources, knowing the responses of interest is important to adequately choose the inputs to include in the model and from which to collect data.

3. Based on the ANOVA method, AT and ST was found to be significant parameters to the outputs of the model. By using mutual information index, the MI measure allows the modelers to go one step further by calculating the proportion of importance of the parameters to the outputs of the model.

To investigate the results more in depth, it was also analyzed how the results of the contingency analysis would change based on the number of bins. Although the proportion of importance of the inputs slightly change based on the number of bins, the conclusion about the most important input did not change for every response of interest and every MI estimation method being investigated. Therefore, in this case, number of bins does not appear to impact the results about which input provides significantly more information to the output. The importance of AT and ST to NIS and TIS does vary based on the number of bins.

Table 64. Results of the contingency analysis to test whether one input provides significantly more information to the output than the other for α-level $= \mathbf{0.05}$.

Method	With or without normalization	NIS Most important input	NIS p-value	TIS Most important input	TIS p-value
Discrete	Without normalization	AT	<0.0001	ST	<0.0001
	With normalization	AT	<0.0001	ST	<0.0001
PDF	Without normalization	AT	<0.0001	AT	<0.0001

With normalization	AT	<0.0001	AT	<0.0001

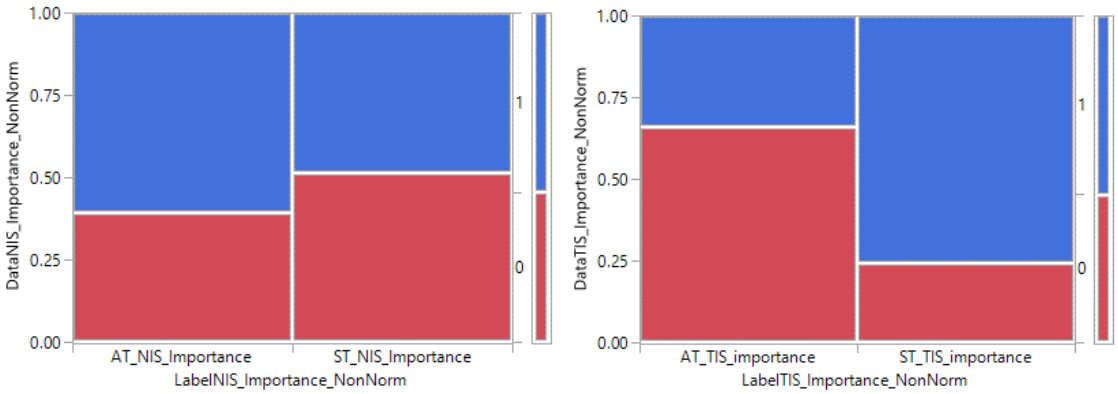

Figure 86. Mosaic plot of importance of inputs for the outputs using discrete empirical distribution to calculate the mutual information measures. Left-side: Importance of inputs for the output $\widehat{\boldsymbol{Y}}_{\mathbf{1}}$ - NIS. Right-side: Importance of inputs for the output $\widehat{\boldsymbol{Y}}_{\mathbf{2}}$ - TIS.

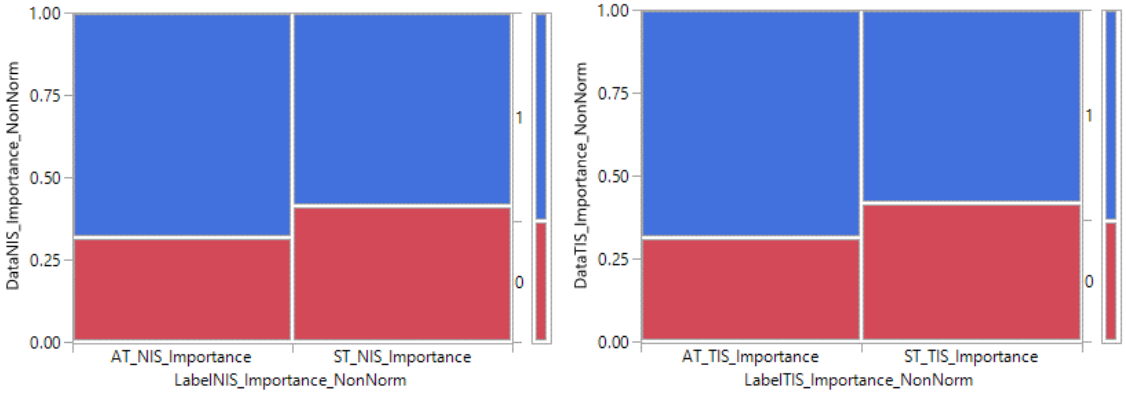

Figure 87. Mosaic plot of importance of inputs for the outputs using probability density distribution to calculate the mutual information measures. Left-side: Importance of inputs for the output $\widehat{\boldsymbol{Y}}_{\mathbf{1}}$ - NIS. Right-side: Importance of inputs for the output $\widehat{\boldsymbol{Y}}_{\mathbf{2}}$ - TIS.

From Table 65, one can see that in general the seed, number of replications, queue model, and traffic intensity impact the importance of the inputs to the outputs. An exception is the factor queue model for the importance of AT and ST to the output NIS

when using discrete empirical distribution. Again, when comparing these results with the ANOVA results, the discrete empirical distribution appears to have better results. According to the ANOVA, seed of interarrival time, traffic intensity, and number of replications were among the factors that were significant for the responses NIS and TIS. Queue model and seed service time were not found to be significant.

Table 65. Results of the contingency analysis to test whether the useful information provided by the input varies based on different factors for α-level $= \mathbf{0.05}$.

Method	With or without normalization	Factors	NIS		TIS	
			Importance of AT p-value	Importance of ST p-value	Importance of AT p-value	Importance of ST p-value
Discrete	Without normalization	Seed interarrival time	<0.0001	<0.0001	<0.0001	<0.0001
		Seed service time	<0.0001	0.0029	0.0401	0.0764
		Number of replications	<0.0001	<0.0001	<0.0001	<0.0001
		Queue model	0.3350	0.2780	<0.0001	<0.0001
		Traffic intensity	<0.0001	<0.0001	<0.0001	<0.0001
	With normalization	Seed interarrival time	<0.0001	<0.0001	<0.0001	<0.0001
		Seed service time	<0.0001	0.0029	0.0401	0.0764
		Number of replications	<0.0001	<0.0001	<0.0001	<0.0001
		Queue model	0.3350	0.2780	<0.0001	<0.0001
		Traffic intensity	<0.0001	<0.0001	<0.0001	<0.0001
PDF	Without normalization	Seed interarrival time	<0.0001	<0.0001	<0.0001	<0.0001
		Seed service time	<0.0001	<0.0001	<0.0001	<0.0001
		Number of replications	<0.0001	<0.0001	<0.0001	<0.0001

	Queue model	0.0021	<0.0001	0.0302	<0.0001
	Traffic intensity	<0.0001	<0.0001	<0.0001	<0.0001
	Seed interarrival time	<0.0001	<0.0001	<0.0001	<0.0001
	Seed service time	<0.0001	<0.0001	<0.0001	<0.0001
With normalization	Number of replications	<0.0001	<0.0001	<0.0001	<0.0001
	Queue model	0.0021	<0.0001	0.0302	<0.0001
	Traffic intensity	<0.0001	<0.0001	<0.0001	<0.0001

4.5. Concluding remarks

In this work, three different applications of entropy measures and mutual information were investigated to support experiment planning and input parameter selection in simulation models. For this, a total of 1620 experiments were run in Simio® University Enterprise Edition v 12.205 with different experiment configurations.

In section 4.4.3, entropy measures were used to support the identification of the seed that leads to the largest uncertainty, if any, in simulation models. For that, regression analysis and Tukey-Kramer multiple comparison test were used. As the results indicate, the entropy measures were able to detect similar statistical differences among the seeds as the SEM, but sometimes the entropy measures were more sensitive and detected more statistical differences than the SEM. However, the entropy measures were not consistent with the SEM method in detecting the seed that led to the largest uncertainty. In this case, there are three different hypotheses: (i) the entropy measure is not accurate, (ii) the SEM is not accurate, or (iii) none of the methods are accurate. When the entropy measures and

the SEM were investigated to detect the seed that led to the largest number of errors, based on the confidence interval resulting from the simulation model and the theoretical value of the input or output, none of the methods appeared to be consistent. However, depending on the parameter values selected, the entropy measure was able to consistently detect the seed that led to the largest number of errors. This highlighted the impact of number of bins on the results of entropy measures.

In section 4.4.4, entropy measures were used to support the identification of the number of replications that leads to the largest uncertainty in simulation models. Similar to section 4.4.3, regression analysis and Tukey-Kramer multiple comparison test was used for this purpose. Considering the normalized entropy measures, the entropy measures estimated either using discrete empirical distribution or probability density distribution showed results similar to the SEM. More specifically, it appears that if one is constrained by time or computer resources, 800 replications is a better choice than 1,000 replications.

In section 4.4.5, mutual information was used to identify the most important parameters to the models based on the response of interest. For that, ANOVA, contingency analysis, and an adaptation of mutual information index were used. The adaptation of mutual information index was proposed so that the measure would add to 1. As the results indicate, the measure appears to be capable to detect the importance of the inputs to the different outputs. However, the results were different based on the estimator being used (discrete vs. PDF). Based on results from section 2.4, section 4.4.3, and also based on the fact that according to Law (2007), time in system appears to be more impacted by service time than by arrival time, it is believed that the discrete empirical distribution is a better

estimate for entropy measures in the context of uncertainty quantification in simulation models. These results also indicate the importance of choosing the responses of interest when modeling a system, as the most significant parameters in a model may be different depending on the response of interest.

According to the experimental results, a few important observations can be made:

1. In general, entropy measures and mutual information are able to detect statistical differences in the groups of seeds used, number of replications, or input parameters similar to other well-known methods in the literature, but the measures are not always able to detect the group that leads to the largest uncertainty consistently with other methods.
2. Normalization of entropy measures is suggested in the literature to minimize the effect of the number of bins in calculating entropy measures. However, as the experimental results suggest, this is more beneficial when the entropy measures calculated with different number of bins are being compared among themselves. Normalization is also required when entropy measures are being used to identify the number of replications that leads to the largest uncertainty. In other situations, normalization may not be required as the results with or without normalization are the same.
3. Although the discrete empirical distribution estimate is not the theoretically correct approach, the estimate appeared to have better results in capturing the deterministic behavior of the simulated queue system and the uncertainty of simulation models overall.

4. The NIS response appears to be more sensitive to seeds than the TIS and the inputs appear to be more sensitive to the number of replications than the outputs.

As future research, it is suggested:

1. To investigate the use of conditional entropy, instead of entropy, as a better measure of uncertainty for simulation models to identify the group of seeds and the number of replications that leads to the largest uncertainty. As previously discussed, the entropy is the average total uncertainty of the input or output, but it also contains the amount of information contained in the input that helps predict the output (or vice-versa). Therefore, a correlation is expected with measures of error, but not necessarily a match. The conditional entropy could be a better measure for these cases where one wants to eliminate the amount of information provided by other variables.
2. To investigate the use of entropy per replication or gain in entropy per replication, $(H(X_{set2}) - H(X_{set1})/(nreps_{set2} - nreps_{set1}))$, as a better measure to identify the number of replications that leads to the largest uncertainty.
3. To investigate the optimum number of bins to be used to calculate entropy measures in the context of simulation models. For this, one could explore optimization models to search for the optimum number of bins, where the objective function is to have the greatest consistency (or minimize the

difference) between the seed that leads to the largest uncertainty identified by the entropy measures and the seed identified by the largest number of errors, such as shown in Equation 83.

$$min \sum_i \left(if\ seed\ H_{b_i} = seed\ E_{b_i} \Rightarrow 0, \qquad otherwise\ \ 1 \right)$$
$$subject\ to$$
$$b_i = 1\ or\ 0\ , i = 1,2, \dots ,10$$
$$\sum_i b_i = 1$$

Equation 83

Where b_i is i^{th} number of clusters.

4. To further explore the optimum number of replications to be used in simulation models with the support of entropy measures by using optimization models. In this case, one can maximize the MI provided by the inputs to the outputs, such as shown in Equation 84.

$$max \sum_j (I(X_1; Y) + I(X_2; Y))$$
$$subject\ to$$
$$r_j = 1\ or\ 0\ , j = 1,2, \dots ,10$$
$$\sum_j r_j = 1$$

Equation 84

Where r_j is j^{th} number of replications.

As previously highlighted, entropy measures are dependent on the number of bins used to calculate the measures. Therefore, adequately defining the number of bins should be a topic of further investigation.

The fact that the method was not validated theoretically is the main limitation of this work. However, in an attempt to mitigate this limitation, the method was compared

with results from other well-known methods from the literature. There are still many open questions about this topic and a vast area of research for further investigation. From the experimental results, although more work needs to be done, the method appears to be capable to support experimental planning and parameter selection in simulation models.

5. CONCLUSIONS

Through empirical results from different simulation experiments on a queue system, this dissertation provided insights about the implications of uncertainty in the results of simulation models and it investigated entropy measures as a method for uncertainty quantification in simulation models.

The first contribution of this dissertation was to discuss the trade-off between model uncertainty and data uncertainty and to show the implications of uncertainty in the results of simulation models through an illustrative queue example. In section 1.1.2, simulation experiments revealed that the accuracy of the model does not necessarily increase with the increase in the number of replications. Using this result, the importance of assessing whether running more replications is economically and computationally attractive was discussed.

In section 2, Shannon's entropy and mutual information measures were investigated as measures of uncertainty in simulation models when using histogram-based method with varying number of bins and normalization methods. The first contribution of this section was to discuss the challenges encountered when computing entropy measures for continuous variables. It was shown that using the existing discrete scheme of correction for continuous variables while estimating entropy measures can still lead to negative values. A similar issue occurs when one adopts $m(x) = E\left[\hat{f}(x)\right]$ in the solution proposed by Jaynes (1957). To overcome these challenges, a data normalization procedure was proposed that minimizes the issues of interpretability of entropy in the continuous case.

The proposed procedure can be implemented when using fixed number of bins, but it cannot be implemented when using formulas that calculate the optimum number of bins. In section 2, it was also shown that when entropy and MI measures were estimated using the histogram-based method with probability density function, the measures tended to decrease with the increase in the number of bins (or decrease in the binwidth) for the same number of replications, which contradicts what is mentioned in the literature. When the measures were estimated using the histogram-based method with discrete empirical distribution, the measures tended to increase with the increase in the number of bins. This led to the hypothesis that in the literature, researchers have been using the discrete empirical distribution to estimate the measures even though the variables may be continuous. To eliminate the effect of the bins on the MI, different normalization methods were assessed and the theoretical normalization of the MI was the method that showed the best results for not altering the behavior of the MI with respect to the number of replications. As it was later shown in section 2, either the probability density function or the discrete empirical distribution was able to capture the uncertainty in the simulation models with respect to different traffic intensities, different seeds, and different parameters values. However, only the discrete empirical distribution was able to adequately capture the deterministic behavior of systems or input parameters like CONWIP or travel time regardless of the number of bins. Using the probability density function, the entropy and MI measures were only able to capture the deterministic behavior for number of bins exceeding 1,000.

Section 3 was an extension of section 2, where Shannon's entropy and mutual information were investigated as measures of uncertainty in simulation models using different estimators, namely kernel-based method, k-nearest neighbors, and fuzzy-histogram-based method. The first contribution of section 3 was to propose a function $m(x) = \hat{f}(x)\big(1 + \hat{f}(x)\big)$ to be used in the approach proposed by Jaynes (1957) to handle the challenges in calculating entropy measures for continuous variables. The function allowed for the calculation of the entropy measures regardless of the choice of the bin and the choice of the density estimation method. Next, a challenge encountered when calculating MI using the product kernel multivariate formula was discussed and a solution was proposed based on discussions found in Silverman (1986). Similar to section 2, the results showed that entropy and MI measures estimated using the kernel-based method, k-nearest neighbors, and fuzzy-histogram-based method were able to capture the uncertainty in the simulation models with respect to different traffic intensities, different seeds, and different parameters values, but were not able to adequately capture the deterministic behavior of systems or input parameters like CONWIP or travel time regardless of the parameter used (such as bandwidth, k-nearest neighbors, or number of bins). The reasons for the estimators not being able to adequately capture the deterministic behavior were discussed in the chapter, which can be used in future research to propose adaptation of the methods to capture the deterministic behavior. For the KNN method and the fuzzy-histogram based method, the entropy was able to capture the deterministic behavior regardless of the choice of the k-nearest neighbors or number of bins. Nevertheless, the MI was not able to capture the deterministic behavior in any of the methods, that is, when

estimated either using the kernel-based method, the KNN method, or the fuzzy-histogram based method. For the kernel-based method, the entropy measures also had issues in capturing the deterministic behavior. In section 3.4.9, an overall comparison in terms of characteristics and results of all estimators was provided.

Finally, section 4 was a continuation of section 2 and section 3, where the ability of entropy measures to quantify uncertainty in simulation models was investigated through a series of applications of the measures for input parameter selection and experiment planning. These applications involved calculating Shannon's entropy and mutual information for full-factorial queue simulation model experiments using stationary univariate distributions. The entropy measures were estimated using histogram-based method with probability density function and the entropy normalization method proposed in section 3.4, as well as using histogram-based method with empirical discrete distribution. Standard error of the mean, regression analysis, and Tukey-Kramer multiple comparison test were used to investigate whether entropy measures could support the identification of the number of replications and the seed that led to the largest uncertainty in the simulation model. ANOVA and contingency analysis were used to investigate whether MI measures could support the identification of the most important input parameters for the simulation model. An adaptation of the mutual information index was proposed to identify the most important input parameters of the simulation model. As the results indicated, the entropy measures were able to detect similar statistical differences among the seeds as the SEM but were not consistent with the SEM method in detecting the seed that led to the largest uncertainty and only the normalized entropy measures

showed results similar to the SEM for identifying the number of replications that leads to the largest uncertainty in simulation models. With respect to detecting the importance of the inputs to the different outputs, the results were different based on the estimator being used (discrete vs. PDF). Based on results from section 2.4, section 4.4.3, and also found by Law (2007), time in system appears to be more impacted by service time than by arrival time, it was experimentally found that the discrete empirical distribution is a better estimate for entropy measures in the context of uncertainty quantification in simulation models. These results also indicated the importance of choosing the response parameters of interest when modeling a system, as the most significant parameters in a model may be different depending on the response parameters of interest. The advantage of the mutual information index over ANOVA is that while ANOVA only allows to indicate whether the input parameter is statistically significant or not for the output, the proposed adaptation of the mutual information index allows to quantify which input parameter is more significant than the other by measuring the importance in terms of percentage.

The main limitations of this dissertation are: (i) only a queue-system was used as an example to run the experiments; (ii) although the entropy and MI measures were compared with other well-known measures reported in the scientific community, the method was not validated theoretically; and (iii) a quantitative way to rank the entropy estimators among themselves was not developed.

As future research, it is recommended: (i) to run similar experiments in different type of systems and to investigate how the responses would change (e.g., flow system, infection-transmission system, etc.); (ii) to propose a framework to validate the work

theoretically; (iii) to propose adaptations for the entropy estimators to adequately measure deterministic behavior; (iv) to investigate the use of conditional entropy, instead of entropy, as a better measure of uncertainty for simulation models to identify the group of seeds and the number of replications that lead to the largest uncertainty; (v) to investigate the use of entropy per replication or gain in entropy per replication as a better measure to identify the number of replications that leads to the largest uncertainty; (vi) to investigate the optimum number of bins to be used to calculate entropy measures in the context of simulation models through optimization models; and, (vii) to investigate the optimum number of replications to be used in simulation models through entropy measures and optimization models. There are a number of open questions about this topic and a vast area of research questions for further investigation. From the experimental results shown in this dissertation, the method appears to be capable of supporting uncertainty quantification, experimental planning, and parameter selection in simulation models.

Based on the results, the recommendation while using the method is to calculate the entropy and MI measures using the histogram-based method with discrete empirical distribution, which although is not the correct estimator given the continuous nature of the variables. This was the estimator that showed the most promising results in detecting the uncertainty in the simulation models and also identifying the most important input parameters to impact the outputs. Moreover, the histogram-based method with discrete empirical distribution estimator did not exhibit challenges with respect to its calculation. If one decides to consider other options, the next recommendation is to follow the

approach proposed by Jaynes (1962) but to use $m(x) = \hat{f}(x)\left(1 + \hat{f}(x)\right)$ as proposed in section 3.4.1.

As final recommendations of this work, there are a few suggestions on how entropy measures can be applied by simulation modelers and other practitioners of the field. Entropy and mutual information measures have the advantage of providing a single number or score for comparison among different scenarios, models, inputs, etc., which is an advantage to be explored in future applications.

Based on Figure 5, it is known that a simulation model contains multiple sources of uncertainty, such as: parameter uncertainty, experimental or output uncertainty, simulation generated inputs, and so on. Therefore, entropy measures can be applied in different contexts to measure each one of these sources, as discussed below. Some possible fields of applications are also discussed.

With respect to parameter uncertainty, mutual information measures can be used to quantify the impact of the inputs on the outputs and to identify from a large number of parameters which ones are the most relevant to be included in the model. Such an application was discussed in section 4. Entropy of the input and conditional entropy of the input given the output can also be used to quantify the uncertainty of the collected data and the uncertainty of the simulation generated data. With respect to the collected data, one can compare the data among themselves and their uncertainty to assess which inputs would bring more uncertainty into the system. With respect to the simulation generated data, one can also use entropy measures to investigate the quality of the seeds of the pseudo random number generator, as discussed in section 4.

With respect to experimental or output uncertainty, one can use the entropy of the output or conditional entropy of the output given the input to quantify the uncertainty of the outputs and determine the best number of replications to run based on the responses of interest and the available computational resources, similar to one of the applications discussed in section 4. Another important application of entropy is as an extension of model verification and validation. As a single measure value, one can also use entropy: (i) to compare multiple models with different input parameters among themselves, for instance, and choose the one that has the lowest uncertainty; and (ii) to validate a model by comparing the uncertainty of the model with the uncertainty of the historical collected data.

Entropy measures can only be calculated after some data is either collected in the real-world or generated in the simulation model. However, after some data is collected and the entropy is calculated, it is possible to aggregate the entropy measures in real-time to the simulation models reports or to the visualization charts. These reports or visualizations can be used for different comparisons as previously discussed in this section. Bootstrap techniques can be used to generate confidence intervals of the entropy measures as well, instead of using a single point estimate.

In terms of field of applications, entropy measures can be applied to a range of simulation models and with different goals, such as: to quantify the variability in the performance of manufacturing design or manufacturing processes; to quantify predictive uncertainty in hydrological models or to improve weather forecasting models like hurricane models; to quantify uncertainty and decision risk in environmental crisis

management models; to quantify risk in financial models, sport models, and medical diagnoses; to select models in image detection or natural language processing; and several other applications in different domains.

As examples of how entropy measures can be used by practitioners in their real-world applications, it is possible to cite: (i) public heatlh practitioners can potentially use mutual information in infectious disease spread models to rank the importance of input parameters such as, periodic (such as daily) vector to human infect rate, periodic human to vector infect rate, periodic human latent rate, periodic human recovery rate, periodic vector mortality rate, percentage of asymptotic cases, periodic transmission rate, among others, on output parameters, such as total duration of the epidemic, number of epidemic waves, total number of exposed individuals, total number of infected individuals, total number of recovered individuals; (ii) public health practitioners can potentially use entropy measures to create a dashboard for analysis and comparison of the different uncertainty values resulting from input parameters collected in different regions, such as daily transmission rate in one country or region versus daily transmission rate in another country or region, or parameters that are difficult to quantify (quantified subjectively), such as the probability that individuals in a population that adhere to control measures; (iii) emergency responders can potentially use entropy measures to evaluate the risks of a wildfire spread given the evolving conditions (e.g., specific wind conditions, weather, number of firefighters dispatched, equipment usage, etc.); (iv) emergency responders can potentially use mutual information to assess how temporal weather data, topographic data, and real spatial fuels data impact the wildfire spread outcome; (v) emergency response

planners can potentially use entropy measures as a single value measure for comparison of the risks involved in the different possible human behaviors displayed during a disaster and, subsequently use mutual information to classify these different human behaviors in different categories, such as low risk, medium risk, and high risk, based on the expected number of total casualties resulting from disaster evacuation simulations; (vi) emergency response planners can potentially use mutual information in hurricane simulations to calculate the impact of air-sea temperatures, mileposts, and terrain on storm intensity, size, and speed, and as these numbers are updated and the storm evolves, dashboards can be developed to show to the user how the uncertainty and the risk of the storm changes from the initial formation until the moment the storm makes landfall; and, (vii) manufacturing and reliability engineers can use reliability system simulation to quantify the uncertainty of producing good parts and equipment breakdown, which, in turn, translates to the risk of producing defective parts and process disruptions respectively.

REFERENCES

Abásolo, D., Escudero, J., Hornero, R., Gómez, C., & Espino, P. (2008). Approximate entropy and auto mutual information analysis of the electroencephalogram in Alzheimer's disease patients. *Medical & biological engineering & computing, 46*(10), 1019-1028.

Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *Bmj, 331*(7521), 903.

Amigo, J. M., Keller, K., & Unakafova, V. A. (2015). On entropy, entropy-like quantities, and applications. *Discrete & Continuous Dynamical Systems- Series B, 20*(10), 3301-3343.

Ankenman, B. E., & Nelson, B. L. (2012). *A quick assessment of input uncertainty*. Paper presented at the Proceedings of the Winter Simulation Conference.

Arefi, M., Viertl, R., & Taheri, S. M. (2012). Fuzzy density estimation. *Metrika, 75*, 5-22.

Asl, M. B., Khalilzadeh, A., Youshanlouei, H. R., & Mood, M. M. (2012). Identifying and ranking the effective factors on selecting Enterprise Resource Planning (ERP) system using the combined Delphi and Shannon Entropy approach. *Procedia-Social and Behavioral Sciences, 41*, 513-520.

Attaran, M., & Zwick, M. (1987). Entropy and other measures of industrial diversification. *Quarterly Journal of Business and Economics*, 17-34.

Awad, A. M., & Alawneh, A. J. (1987). Application of Entropy to a life-time model. *IMA Journal of Mathematical Control and Information, 4*, 143-147.

Bao-Gang, H., & Yong, W. (2008). Evaluation criteria based on mutual information for classifications including rejected class. *Acta Automatica Sinica, 34*(11), 1396-1403.

Barde, M. P., & Barde, P. J. (2012). What to use to express the variability of data: Standard deviation or standard error of mean? *Perspectives in clinical research, 3*(3), 113.

Barton, R. R. (2012). *Tutorial: input uncertainty in outout analysis.* Paper presented at the Proceedings of the 2012 Winter Simulation Conference.

Barton, R. R., Nelson, B. L., & Xie, W. (2010). *A framework for input uncertainty analysis.* Paper presented at the Proceedings of the 2010 Winter Simulation Conference.

Barton, R. R., Nelson, B. L., & Xie, W. (2014). Quantifying input uncertainty via simulation confidence intervals. *INFORMS Journal on Computing, 26*(1), 74-87. doi:http://dx.doi.org/10.1287/ijoc.2013.0548

Barton, R. R., & Schruben, L. W. (1993). *Uniform and bootstrap resampling of empirical distributions.* Paper presented at the Proceedings of the 25th conference on Winter simulation.

Baudin, M., Dutfoy, A., Iooss, B., & Popelin, A.-L. (2015). OpenTURNS: An industrial software for uncertainty quantification in simulation. In R. Ghanem, D. Higdon, & H. Owhadi (Eds.), *Handbook of Uncertainty Quantification* (pp. 1-38). Cham, Switzerland: Springer.

Baumert, M., Baier, V., Haueisen, J., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Voss, A. (2004). Forecasting of life threatening arrhythmias using the compression entropy of heart rate. *Methods of information in medicine, 43*(02), 202-206.

Beirlant, J., Dudewicz, E. J., Györfi, L., & Van der Meulen, E. C. (1997). Nonparametric entropy estimation: An overview. *International Journal of Mathematical and Statistical Sciences, 6*(1), 17-39.

Berger, A., Della Pietra, S. A., & Della Pietra, V. J. (1996). A maximum entropy approach to natural language processing. *Computational linguistics, 22*(1), 39-71.

Biller, B., & Gunes, C. (2010). *Introduction to simulation input modeling.* Paper presented at the Proceedings of the 2010 Winter Simulation Conference.

Biller, B., & Nelson, B. L. (2002). *Answers to the top ten input modeling questions.* Paper presented at the Proceedings of the 2002 Winter Simulation Conference.

Birgé, L., & Rozenholc, Y. (2006). How many bins should be put in a regular histogram. *ESAIM: Probability and Statistics, 10*, 24-45.

Bisetti, F., Kim, D., Knio, O., Long, Q., & Tempone, R. (2016). Optimal Bayesian Experimental Design for Priors of Compact Support with Application to Shock-Tube Experiments for Combustion Kinetics. *International Journal for Numerical Methods in Engineering, 108*(2), 136-155.

Blanco, D. D., Casini, H., Hung, L.-Y., & Myers, R. C. (2013). Relative entropy and holography. *Journal of High Energy Physics, 2013*(8), 60.

Bonnlander, B. V., & Weigend, A. S. (1994). *Selecting input variables using mutual information and nonparametric density estimation.* Paper presented at the Proceedings of the 1994 International Symposium on Artificial Neural Networks (ISANN'94).

Box, G. E. P., & Meyer, R. D. (1986a). An analysis for unreplicated fractional factorials. *Technometrics, 28*(1), 11-18.

Box, G. E. P., & Meyer, R. D. (1986b). Dispersion effects from fractional designs. *Technometrics, 28*(1), 19-27.

Bravi, A., Longtin, A., & Seely, A. J. E. (2011). Review and classification of variability analysis techniques with clinical applications. *Biomedical engineering online, 10*(90), 1-27.

Brownlee, J. (2019). A Gentle Introduction to Cross-Entropy for Machine Learning. Retrieved from https://machinelearningmastery.com/cross-entropy-for-machine-learning/

Callao, M. P. (2014). Multivariate experimental design in environmental analysis. *Trends in Analytical Chemistry, 62*, 86-92.

Castro, R. (2015). The empirical distribution function and the histogram. *Lecture Notes, 2WS17-Advanced Statistics. Department of Mathematics, Eindhoven University of Technology, 4*.

Cavanaugh, J. E. (1997). Unifying the derivations for the Akaike and corrected Akaike information criteria. *Statistics & Probability Letters, 33*(2), 201-208.

Chaloner, K., & Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, 273-304.

Chandrasekaran, V., & Shah, P. (2017). Relative entropy optimization and its applications. *Mathematical Programming, 161*(1-2), 1-32.
Chen, P., Quarteroni, A., & Rozza, G. (2013). Simulation-based uncertainty quantification of human arterial network hemodynamics. *International Journal for Numerical Methods in Biomedical Engineering, 29*(6), 698-721.
Chen, X., Kar, S., & Ralescu, D. A. (2012). Cross-entropy measure of uncertain variables. *Information Sciences, 201*, 53-60.
Cheng, H.-D., & Chen, Y.-H. (1999). Fuzzy partition of two-dimensional histogram and its application to thresholding. *Pattern Recognition, 32*, 825-843.
Christley, R. M., Mort, M., Wynne, B., Wastling, J. M., Heathwaite, A. L., Pickup, R., . . . Latham, S. M. (2013). "Wrong, but Useful": Negotiating Uncertainty in Infectious Disease Modelling. *Plos One, 8*(10). doi:10.1371/journal.pone.0076277
Cicirelli, F., Furfaro, A., & Nigro, L. (2011). Modelling and simulation of complex manufacturing systems using statechart-based actors. *Simulation Modelling Practice and Theory, 19*(2), 685-703.
Clyde, M. A. (2001). Experimental design: A Bayesian perspective. *International Encyclopia Social and Behavioral Sciences*, 1-22.
Council for Regulatory Environmental Modeling. (2009). Web-based Training on Best Modeling Practices and Technical Modeling Issues: Sensitivity and Uncertainty Analyses.
Critchfield, G. C., & Willard, K. E. (1986). Probabilistic analysis of decision trees using Monte Carlo simulation. *Medical Decision Making, 6*(2), 85-92.
De Boer, P.-T., Kroese, D. P., Mannor, S., & Rubinstein, R. Y. (2005). A tutorial on the cross-entropy method. *Annals of operations research, 134*(1), 19-67.
Dean, A., & Lewis, S. (2006). *Screening: methods for experimentation in industry, drug discovery, and genetics*: Springer Science & Business Media.
DeVolder, B., Glimm, J., Grove, J. W., Kang, Y., Lee, Y., Pao, K., . . . Ye, K. (2002). Uncertainty quantification for multiscale simulations. *Journal of Fluids Engineering-Transactions of the Asme, 124*(1), 29-41. doi:10.1115/1.1445139
Dionisio, A., Menezes, R., & Mendes, D. A. (2004). Mutual information: a measure of dependency for nonlinear time series. *Physica A: Statistical Mechanics and its Applications, 344*(1-2), 326-329. doi:10.1016/j.physa.2004.06.144
Egnal, G., & Daniilidis, K. (2000). Image registration using mutual information. *Technical Reports (CIS)*, 117.
Estévez, P. A., Tesmer, M., Perez, C. A., & Zurada, J. M. (2009). Normalized mutual information feature selection. *IEEE Transactions on Neural Networks, 20*(2), 189-201. doi:10.1109/TNN.2008.2005601
Fajardo, J. A. (2014). A criterion for the fuzzy set estimation of the density function. *Brazilian Journal of Probability and Statistics, 28*(3), 301-312.
Foutz, R. V. (1980). A test for goodness-of-fit based on an empirical probability measure. *The Annals of Statistics*, 989-1001.
Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A, 33*(2), 1134.

Frey, H. C., & Patil, S. R. (2002). Identification and review of sensitivity analysis methods. *Risk analysis, 22*(3), 553-578.
Fuglede, B., & Topsoe, F. (2004). *Jensen-Shannon divergence and Hilbert space embedding.* Paper presented at the International Symposium onInformation Theory, 2004. ISIT 2004. Proceedings.
Garner, D. M., & Ling, B. W.-K. (2014). Measuring and locating zones of chaos and irregularity. *Journal of Systems Science and Complexity, 27*(3), 494-506.
Gautam, N. (2012). *Analysis of queues: methods and applications*: CRC Press.
Ghosh, J. (2002, June, 2002). *Multiclassifier systems: Back to the future.* Paper presented at the International Workshop on Multiple Classifier Systems, Berlin, Heidelberg.
Greven, A., Keller, G., & Warnecke, G. (2014). *Entropy* (Vol. 47): Princeton University Press.
Gutierrez-Osuna, R. (2020). L7: Kernel density estimation. Retrieved from http://research.cs.tamu.edu/prism/lectures/pr/pr_l7.pdf
Haeri, M. A., & Ebadzadeh, M. M. (2014). Estimation of mutual information by the fuzzy histogram. *Fuzzy Optimization and Decision Making, 13*(3), 287-318.
Hanson, K., & Hemez, F. (2003). *Uncertainty quantification of simulation codes based on experimental data.* Paper presented at the Proceedings of the 41st Aerospace Sciences Meeting and Exhibit.
Härdle, W. K., Müller, M., Sperlich, S., & Werwatz, A. (2012). *Nonparametric and semiparametric models*: Springer Science & Business Media.
Haverkamp, S., Srinivasan, R., Frede, H., & Santhi, C. (2002). Subwatershed spatial analysis tool: discretization of a distributed hydrologic model by statistical criteria 1. *JAWRA Journal of the American Water Resources Association, 38*(6), 1723-1733.
Havrda, J., & Charvát, F. (1967). Quantification method of classification processes. Concept of structural $ a $-entropy. *Kybernetika, 3*(1), (30)-35.
He, J., & Kolovos, A. (2018). Bayesian maximum entropy approach and its applications: a review. *Stochastic Environmental Research and Risk Assessment, 32*(4), 859-877.
Hernandez, M., Lane, L. J., Stone, J. J., Martinez, J. G., & Kidwell, M. (1997). *Hydrologic model performance evaluation applying the entropy concept as a function of precipitation network density.* Paper presented at the International Congress on Modelling and Simulation Proceedings, University of Tasmania.
Hill, D. L., Batchelor, P. G., Holden, M., & Hawkes, D. J. (2001). Medical image registration. *Physics in medicine & biology, 46*(3), R1.
Hongqiao, Y., Xihua, L., Fei, W., & Weizi, L. (2009). *Multi-agent based modeling and simulation of complex system in hospital.* Paper presented at the 2009 16th International Conference on Industrial Engineering and Engineering Management.
Hopper, T. (2021). Cross Entropy and KL Divergence. Retrieved from https://tdhopper.com/blog/cross-entropy-and-kl-divergence
Huan, X., & Marzouk, Y. (2014). Gradient-based stochastic optimization methods in Bayesian experimental design. *International Journal for Uncertainty Quantification, 4*(6).

Huang, X. (2012). An entropy method for diversified fuzzy portfolio selection. *International Journal of Fuzzy Systems, 14*(1), 160-165.
Hyndman, R. J. (1995). The problem with sturges rule for constructing histograms.
Jaynes, E. T. (1957). Information theory and statistical mechanics. II. *Physical review, 108*(2), 171.
Jaynes, E. T. (1962). Information theory and statistical mechanics. In *Statistical Physics* (Vol. 3, pp. 252). New York: Brandeis Summer Institute.
Jaynes, E. T. (1968). Prior probabilities. *IEEE Transactions on Systems Science and Cybernetics, 4*(3), 227-241.
Kanazawa, Y. (1993). Hellinger distance and Akaike's information criterion for the histogram. *Statistics & probability letters, 17*(4), 293-298.
Kannathal, N., Choo, M. L., Acharya, U. R., & Sadasivan, P. K. (2005). Entropies for detection of epilepsy in EEG. *Computer methods and programs in biomedicine, 80*(3), 187-194.
Kapur, J. N. (1983). A Comparative Assessment of Various Measures of Entropy. *Journal of Information and Optimization Sciences, 4*(3), 207-232. doi:10.1080/02522667.1983.10698762
Kapur, J. N., & Kesavan, H. K. (1990). The inverse MaxEnt and MinxEnt principles and their applications. In *Maximum Entropy and Bayesian Methods* (pp. 433-450): Springer.
Katok, A. (2007). Fifty years of entropy in dynamics: 1958--2007. *Journal of Modern Dynamics, 1*(4), 545.
Kinney, J. B., & Atwal, G. S. (2014). Equitability, mutual information, and the maximal information coefficient. *Proceedings of the National Academy of Sciences, 111*(9), 3354-3359. doi:10.1073/pnas.1309933111
Kitching, R. P., Hutber, A. M., & Thrusfield, M. V. (2005). A review of foot-and-mouth disease with special consideration for the clinical and epidemiological factors relevant to predictive modelling of the disease. *Veterinary Journal, 169*(2), 197-209. doi:10.1016/j.tvjl.2004.06.001
Kittaneh, O. A., Khan, M. A., Akbar, M., & Bayoud, H. A. (2016). Average entropy: a new uncertainty measure with application to image segmentation. *The American Statistician, 70*(1), 18-24.
Koshkin, G. M. (2014). Smooth estimators of the reliability functions for non-restorable elements. *Russian Physics Journal, 57*(5), 672-681.
Kowalski, A. M., Martin, M. T., Plastino, A., & Judge, G. (2012). On extracting probability distribution information from time series. *Entropy, 14*(10), 1829-1841.
Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E, 69*(6), 066138.
Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics, 22*(1), 79-86.
Lacasa, L., & Just, W. (2017). Visibility graphs and symbolic dynamics. *arXiv preprint arXiv:1704.06467*.
Law, A. M. (2007). *Simulation Modeling and Analysis* (4th ed.). New York, NY: McGraw-Hill Higher Education.

Legg, P. A., Rosin, P. L., Marshall, D., & Morgan, J. E. (2013). Improving accuracy and efficiency of mutual information for multi-modal retinal image registration using adaptive probability density estimation. *Computerized Medical Imaging and Graphics, 37*(7-8), 597-606.

Li, Q., & Racine, J. S. (2007). *Nonparametric econometrics: theory and practice*. New Jersey: Princeton University Press.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 986-1005.

Loquin, K., & Strauss, O. (2006). Fuzzy histograms and density estimation. In *Soft Methods in Probability and Statistics* (pp. 45-52). Bristol: Springer.

Loquin, K., & Strauss, O. (2008). Histogram density estimators based upon a fuzzy partition. *Statistics and Probability Letters, 78*, 1863-1868.

Lotfi, F. H., & Fallahnejad, R. (2010). Imprecise Shannon's entropy and multi attribute decision making. *Entropy, 12*(1), 53-62.

Malakar, N. K., & Knuth, K. H. (2011). *Entropy-Based Search Algorithm for Experimental Design.* Paper presented at the AIP Conference Proceedings.

Marelli, S., & Sudret, B. (2014). UQLab: A framework for uncertainty quantification in Matlab. In *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management* (pp. 2554-2563).

Massey, J. (1990). *Causality, feedback and directed information.* Paper presented at the Proc. Int. Symp. Inf. Theory Applic.(ISITA-90).

McDaid, A. F., Greene, D., & Hurley, N. (2011). Normalized mutual information to evaluate overlapping community finding algorithms. *arXiv preprint arXiv:1110.2515*, 1-3.

Milnor, J. (1988). On the entropy geometry of cellular automata. *Complex Systems, 2*(3), 357-385.

Moddemeijer, R. (1989). On estimation of entropy and mutual information of continuous distributions. *Signal processing, 16*(3), 233-248.

Moon, Y.-I., Rajagopalan, B., & Lall, U. (1995). Estimation of mutual information using kernel density estimators. *Physical Review E, 52*(3), 2318.

Mousavian, Z., Kavousi, K., & Masoudi-Nejad, A. (2016). Information theory in systems biology. Part I: Gene regulatory and metabolic networks. *Seminars in cell & developmental biology, 51*, 3-13.

Mu, X.-Z., & Hu, G.-W. (2018). Analysis of Venezuela's oil-oriented economy: from the perspective of entropy. *Petroleum Science, 15*(1), 200-209.

Nelson, B. L. (1987a). A perspective on variance reduction in dynamic simulation experiments. *Communications in Statistics - Simulation and Computation, 16*(2), 385-426.

Nelson, B. L. (1987b). *Variance reduction for simulation practitioners.* Paper presented at the Proceedings of the 1987 Winter Simulation conference.

Nigam, K., Lafferty, J., & McCallum, A. (1999). *Using maximum entropy for text classification.* Paper presented at the IJCAI-99 workshop on machine learning for information filtering.

Oberkampf, W. L., DeLand, S. M., Rutherford, B. M., Diegert, K. V., & Alvin, K. F. (2002). Error and uncertainty in modeling and simulation. *Reliability Engineering & System Safety, 75*(3), 333-357.

Orozco-Arroyave, J. R., Arias-Londono, J. D., Vargas-Bonilla, J. F., & Nöth, E. (2013). *Analysis of speech from people with Parkinson's disease through nonlinear dynamics.* Paper presented at the International Conference on Nonlinear Speech Processing.

Pace, R. K. (1995). Parametric, semiparametric, and nonparametric estimation of characteristic values within mass assessment and hedonic pricing models. *The Journal of Real Estate Finance and Economics, 11*(3), 195-217.

Pham, T. D. (2013). *KS entropy of images.* Paper presented at the 2013 20th International Conference on Systems, Signals and Image Processing (IWSSIP).

Pham, T. D. (2016). The Kolmogorov–Sinai entropy in the setting of fuzzy sets for image texture analysis and classification. *Pattern Recognition, 53*, 229-237. doi:https://doi.org/10.1016/j.patcog.2015.12.012

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling, 190*(3-4), 231-259.

Phillips, S. J., Dudík, M., & Schapire, R. E. (2004). *A maximum entropy approach to species distribution modeling.* Paper presented at the Proceedings of the twenty-first international conference on Machine learning.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences, 88*(6), 2297-2301.

Principe, J. C., Xu, D., Zhao, Q., & Fisher, J. W. (2000). Learning from examples with information theoretic criteria. *Journal of VLSI signal processing systems for signal, image and video technology, 26*(1-2), 61-77. doi:10.1023/A:1008143417156

Qin, Z., Li, X., & Ji, X. (2009). Portfolio selection based on fuzzy cross-entropy. *Journal of Computational and Applied mathematics, 228*(1), 139-149.

Rao, M., Chen, Y., Vemuri, B. C., & Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE transactions on Information Theory, 50*(6), 1220-1228.

Rényi, A. (1961). *On measures of entropy and information.* Paper presented at the Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics.

Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., . . . Sabeti, P. C. (2011). Detecting novel associations in large data sets. *science, 334*(6062), 1518-1524.

Reza, F. M. (1961). *An introduction to information theory*: McGraw Hill.

Rhea, C. K., Silver, T. A., Hong, S. L., Ryu, J. H., Studenka, B. E., Hughes, C. M. L., & Haddad, J. M. (2011). Noise and Complexity in Human Postural Control: Interpreting the Different Estimations of Entropy. *Plos One, 6*(3). doi:ARTN e1769610.1371/journal.pone.0017696

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology, 278*(6), H2039-H2049.

Rickles, D., Hawe, P., & Shiell, A. (2007). A simple guide to chaos and complexity. *Journal of Epidemiology & Community Health, 61*(11), 933-937.

Rossi, F., Lendasse, A., François, D., Wertz, V., & Verleysen, M. (2006). Mutual information for the selection of relevant variables in spectrometric nonlinear modelling. *Chemometrics and intelligent laboratory systems, 80*(2), 215-226.

Roy, C. J., & Oberkampf, W. L. (2011). A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering, 200*(25-28), 2131-2144.

Rubinstein, R. Y. (2001). Combinatorial optimization, cross-entropy, ants and rare events. In *Stochastic optimization: algorithms and applications* (pp. 303-363): Springer.

Ryan, E. G. (2014). *Contributions to Bayesian experimental design.* Queensland University of Technology,

Ryan, E. G., Drovandi, C. C., Thompson, M. H., & Pettitt, A. N. (2014). Towards Bayesian experimental design for nonlinear models that require a large number of sampling times. *Computational Statistics and Data Analysis, 70*, 45-60.

Şahin, R. (2017). Cross-entropy measure on interval neutrosophic sets and its applications in multicriteria decision making. *Neural Computing and Applications, 28*(5), 1177-1187.

Schefzik, R., Thorarinsdottir, T. L., & Gneiting, T. (2013). Uncertainty quantification in complex simulation models using ensemble copula coupling. *Statistical Science, 28*(4), 616-640.

Scheidegger, A. P. G., Banerjee, A., & Pereira, T. F. (2018). *Uncertainty quantification in simulation models: a proposed framework and application through case study.* Paper presented at the 2018 Winter Simulation Conference (WSC).

Schreiber, T. (2000). Measuring information transfer. *Physical review letters, 85*(2), 461-465.

Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika, 66*(3), 605-610.

Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*: John Wiley & Sons.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*, 379-423.

Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell system technical journal, 30*(1), 50-64.

Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review (Reprinted with corrections from The Bell System Technical Journal), 5*(1), 3-55.

Shore, J., & Johnson, R. (1981). Properties of cross-entropy minimization. *IEEE Transactions on Information theory, 27*(4), 472-482.

Silva, F. M. H. S. P. d., Silva Filho, A. C., Crescêncio, J. C., & Gallo, L. (2012). *Applying Lyapunov exponents in heart rate time series to identify the anaerobic threshold in healthy men.* Paper presented at the Computing in Cardiology (CinC), 2012.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis* (Vol. 26): CRC press.
Song, E., & Nelson, B. L. (2013). *A quicker assessment of input uncertainty.* Paper presented at the Proceedings of the 2013 Winter Simulation Conference.
Song, E., & Nelson, B. L. (2015). Quickly assessing contributions to input uncertainty. *IIE Transactions, 47*(9), 893-909.
Song, E., Nelson, B. L., & Pegden, C. D. (2014). *Advanced tutorial: Input uncertainty quantification.* Paper presented at the Proceedings of the 2014 Winter Simulation Conference.
Steuer, R., Kurths, J., Daub, C. O., Weise, J., & Selbig, J. (2002). The mutual information: detecting and evaluating dependencies between variables. *Bioinformatics, 18*(suppl_2), S231-S240.
Stone, J. V. (2015). *Information theory: A tutorial introduction* (1st ed.). Lexington, KY: Sebtel Press.
Strehl, A., & Ghosh, J. (2002). Cluster ensembles---a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research, 3*(Dec), 583-617.
Sturges, H. A. (1926). The choice of a class interval. *Journal of the american statistical association, 21*(153), 65-66.
Székely, G. J., & Rizzo, M. L. (2009). Brownian distance covariance. *The annals of applied statistics*, 1236-1265.
Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics, 35*(6), 2769-2794.
Tapia, J. E., & Perez, C. A. (2013). Gender classification based on fusion of different spatial scale features selected by mutual information from histogram of LBP, intensity, and shape. *IEEE transactions on information forensics and security, 8*(3), 488-499.
Telesca, L., Caggiano, R., Lapenna, V., Lovallo, M., Trippetta, S., & Macchiato, M. (2008). The Fisher information measure and Shannon entropy for particulate matter measurements. *Physica A: Statistical Mechanics and its Applications, 387*(16-17), 4387-4392.
Tesmer, M., & Estévez, P. A. (2004). *AMIFS: Adaptive feature selection by using mutual information.* Paper presented at the Proceedings of the 2004 IEEE International Joint Conference on Neural Networks.
Truebner, S., Cygankiewicz, I., Schroeder, R., Baumert, M., Vallverdu, M., Caminal, P., . . . Voss, A. (2006). Compression entropy contributes to risk stratification in patients with cardiomyopathy. *Biomedizinische Technik, 51*(2), 77.
Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics, 52*(1), 479-487.
van Den Berg, J., Curtis, A., & Trampert, J. (2003). Optimal nonlinear Bayesian experimental design: an application to amplitude versus offset experiments. *Geophysical Journal International, 155*(2), 411-421.

Venkatramanan, S., Lewis, B., Chen, J., Higdon, D., Vullikanti, A., & Marathe, M. (2018). Using data-driven agent-based models for forecasting emerging infectious diseases. *Epidemics, 22*, 43-49. doi:https://doi.org/10.1016/j.epidem.2017.02.010

Vinh, N. X., Epps, J., & Bailey, J. (2010). Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *The Journal of Machine Learning Research, 11*, 2837-2854.

Vrieze, S. I. (2012). Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological methods, 17*(2), 228.

Walsh, S. N., Wildey, T. M., & Jakeman, J. D. (2018). Optimal experimental design using a consistent Bayesian approach. *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg, 4*(1).

Wand, M. P. (1997). Data-based choice of histogram bin width. *The American Statistician, 51*(1), 59-64.

Wang, Y., & Hu, B.-G. (2009). *Derivations of normalized mutual information in binary classifications.* Paper presented at the 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery.

Waterman, M. S., & Whiteman, D. E. (1978). Estimation of probability densities by empirical density functions. *International Journal of Mathematical Education in Science and Technology, 9*(2), 127-137.

Wehrl, A. (1978). General properties of entropy. *Reviews of Modern Physics, 50*(2), 221.

Wijaya, D. R., Sarno, R., & Zulaika, E. (2017). Information Quality Ratio as a novel metric for mother wavelet selection. *Chemometrics and intelligent laboratory systems, 160*, 59-71.

Xie, W., Nelson, B. L., & Barton, R. R. (2014a). A Bayesian Framework for Quantifying Uncertainty in Stochastic Simulation. *Operations Research, 62*(6), 1439-1452. doi:10.1287/opre.2014.1316

Xie, W., Nelson, B. L., & Barton, R. R. (2014b). *Statistical Uncertainty Analysis for Stochastic Simulation with Dependent Input Models.* Paper presented at the Proceedings of the 2014 Winter Simulation Conference (Wsc).

Xiong, W., Faes, L., & Ivanov, P. C. (2017). Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Physical Review E, 95*(6), 062114. doi:10.1103/PhysRevE.95.062114

Xu, Y., Ning, X., Chen, Y., & Wang, J. (2004). Mode entropy and dynamical analysis of irregularity for HFECG. *Chinese Science Bulletin, 49*(17), 1886-1890.

Yentes, J. M., Hunt, N., Schmid, K. K., Kaipust, J. P., McGrath, D., & Stergiou, N. (2013). The appropriate use of approximate entropy and sample entropy with short data sets. *Annals of biomedical engineering, 41*(2), 349-365.

Zhang, W., & Wang, J. (2017). Nonlinear stochastic exclusion financial dynamics modeling and complexity behaviors. *Nonlinear Dynamics, 88*(2), 921-935.

Zhang, Z., Sun, S., Yi, M., Wu, X., & Ding, Y. (2015). MIC as an Appropriate Method to Construct the Brain Functional Network. *BioMed Research International, 2015*, 825136. doi:10.1155/2015/825136

APPENDIX A

CONFIGURATION OF EXPERIMENTS AND RESULTS

Table 66. Experiments to assess the quality of entropy and mutual information as measures of uncertainty quantification in simulation models.

Experiment number	Queue model (Kendall's notation)	Number of servers s	Inter-arrival time $1/\lambda$ [minutes]	Inter-arrival time seed	Service time $1/\mu$ [minutes]	Service time seed	Travel time [minutes]	Traffic intensity ρ	Number of replications
1 to 10	M/M/1	1	20	4	2	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
11 to 20					10			0.50	
21 to 30					18.000018			0.90	
31 to 40	M/M/3	3	6.666667	4	2	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
41 to 50					10			0.50	
51 to 60					18.000018			0.90	
61 to 70	M/M/10	10	2	4	2	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
71 to 80					10			0.50	
81 to 90					18.000018			0.90	
91 to 100	M/G/1	1	20	4	$\sigma^2 = .01\frac{1}{\mu}^2$	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
101 to 110			4					0.50	
111 to 120			2.222222					0.90	
121 to 130	M/G/1	1	20	4	$\sigma^2 = .25\frac{1}{\mu}^2$	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
131 to 140			4					0.50	
141 to 150			2.222222					0.90	
151 to 160	G/G/1	1	$\sigma^2 = .01\frac{1}{\lambda}^2$	4	$\sigma^2 = .01\frac{1}{\mu}^2$	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
161 to 170					$\sigma^2 = .01\frac{1}{\mu}^{10}$			0.50	
171 to 180					$\sigma^2 = .01\frac{1}{\mu}^{18.000018}$			0.90	
181 to 190	G/G/1	1	$\sigma^2 = .25\frac{1}{\lambda}^{20}$	4	$\sigma^2 = .01\frac{1}{\mu}^2$	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
191 to 200					$\sigma^2 = .01\frac{1}{\mu}^{10}$				
201 to 210					$8\sigma^2 = .01\frac{1}{\mu}^{18.000018}$				
211 to 220	G/G/1	1	$\sigma^2 = .01\frac{1}{\lambda}^{20}$	4	$\sigma^2 = .25\frac{1}{\mu}^2$	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
221 to 230					$\sigma^2 = .25\frac{1}{\mu}^{10}$			0.50	

231 to 240					18.000018 $\sigma^2 = .25\frac{1}{\mu}$			0.90	
241 to 250					$\sigma^2 = .25\frac{1}{\mu}$			0.10	
251 to 260	G/G/1	1	20 $\sigma^2 = .25\frac{1}{\lambda}$	4	10 $\sigma^2 = .25\frac{1}{\mu}$	6	not applicable	0.50	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
261 to 270					18.000018 $\sigma^2 = .25\frac{1}{\mu}$			0.90	
271 to 280	G/G/3	3	6.666667 $\sigma^2 = .01\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .01\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
281 to 290	G/G/3	3	6.666667 $\sigma^2 = .25\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .01\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
291 to 300	G/G/3	3	6.666667 $\sigma^2 = .01\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .25\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
301 to 310	G/G/3	3	6.666667 $\sigma^2 = .25\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .25\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
311 to 320	G/G/10	10	2 $\sigma^2 = .01\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .01\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
321 to 330	G/G/10	10	2 $\sigma^2 = .25\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .01\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
331 to 340	G/G/10	10	2 $\sigma^2 = .01\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .25\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
341 to 350	G/G/10	10	2 $\sigma^2 = .25\frac{1}{\lambda}$	4	18.000018 $\sigma^2 = .25\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
351 to 360	M/M/1	1	20	8	2	15	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
361 to 370					10			0.50	
371 to 380					18.000018			0.90	
381 to 390	M/M/3	3	6.666667	8	2	15	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
391 to 400					10			0.50	
401 to 410					18.000018			0.90	
411 to 420	M/M/10	10	2	8	2	15	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
421 to 430					10			0.50	
431 to 440					18.000018			0.90	
441 to 450	M/G/1	1	20	8	2	15		0.10	

451 to 460			4		$\sigma^2 = .01\frac{1}{\mu}$		not applicable	0.50	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
461 to 470	G/G/3	3	6.666667 $\sigma^2 = .01\frac{1}{\lambda}$	8	18.000018 $\sigma^2 = .01\frac{1}{\mu}$	15	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
471 to 480	M/M/1	1	20	20	2	5	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
481 to 490					10			0.50	
491 to 500					18.000018			0.90	
501 to 510	M/M/3	3	6.666667	20	2	5	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
511 to 520					10			0.50	
521 to 530					18.000018			0.90	
531 to 540	M/M/10	10	2	20	2	5	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
541 to 550					10			0.50	
551 to 560					18.000018			0.90	
561 to 570	M/G/1	1	20	20	2	5	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
571 to 580			4		$\sigma^2 = .01\frac{1}{\mu}$			0.50	
581 to 590	G/G/3	3	6.666667 $\sigma^2 = .01\frac{1}{\lambda}$	20	18.000018 $\sigma^2 = .01\frac{1}{\mu}$	5	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
591 to 600	M/M/1	1	45	4	4.5	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
601 to 610					25.500002			0.50	
611 to 620					40.5			0.90	
621 to 630	M/M/3	3	15	4	45	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
631 to 640					25.500002			0.50	
641 to 650					40.5			0.90	
651 to 660	M/M/10	10	4.5	4	4.5	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
661 to 670					25.500002			0.50	
671 to 680					40.5			0.90	
681 to 690	M/G/1	1	45	4	4.5	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
691 to 700			9		$\sigma^2 = .01\frac{1}{\mu}$			0.50	
701 to 710	G/G/3	3	15 $\sigma^2 = .01\frac{1}{\lambda}$	4	40.5 $\sigma^2 = .01\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
711 to 720	M/M/1	1	180	4	18	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
721 to 730					90			0.50	
731 to 740					162			0.90	
741 to 750	M/M/3	3	60	4	18	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
751 to 760					90			0.50	
761 to 770					162			0.90	
771 to 780	M/M/10	10	18	4	18	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
781 to 790					90			0.50	
791 to 800					162			0.90	
801 to 810	M/G/1	1	180	4	18	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
811 to 820			36		$\sigma^2 = .01\frac{1}{\mu}$			0.50	
821 to 830	G/G/3	3	60 $\sigma^2 = .01\frac{1}{\lambda}$	4	162 $\sigma^2 = .01\frac{1}{\mu}$	6	not applicable	0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
831 to 840	CONWIP	1	not applicable	not applicable	2	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
841 to 850					10			0.50	
851 to 860					18.000018			0.90	

861 to 870 871 to 880 881 to 890	CONWIP	3	not applicable	not applicable	2 10 18.000018	6	not applicable	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
891 to 900 901 to 910 911 to 920	CONWIP	10	not applicable	not applicable	2 10 18.000018	6	not applicable	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
921 to 930	CONWIP	1	not applicable	not applicable	$\sigma^2 = .01\frac{1}{\mu}^2$	6	not applicable	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
931 to 940 941 to 950 951 to 960	M/M/1	1	20	4	2 10 18.000018	6	Det(10) [1]	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
961 to 970 971 to 980 981 to 990	M/M/3	3	6.666667	4	2 10 18.000018	6	Det(10) [1]	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
991 to 1,000 1,001 to 1,010 1,011 to 1,020	M/M/10	10	2	4	2 10 18.000018	6	Det(10) [1]	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
1,021 to 1,030	M/G/1	1	20	4	$\sigma^2 = .01\frac{1}{\mu}^2$	6	Det(10) [1]	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
1,031 to 1,040 1,041 to 1,050 1,051 to 1,060	M/M/1	1	20	4	2 10 18.000018	6	Exp(10,0) [2]	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
1,061 to 1,070 1,071 to 1,080 1,081 to 1,090	M/M/3	3	6.666667	4	2 10 18.000018	6	Exp(10,0) [2]	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
1,091 to 1,100 1,101 to 1,110 1,111 to 1,120	M/M/10	10	2	4	2 10 18.000018	6	Exp(10,0) [2]	0.10 0.50 0.90	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500
1,121 to 1,130	M/G/1	1	20	4	$\sigma^2 = .01\frac{1}{\mu}^2$	6	Exp(10,0) [2]	0.10	10, 20, 50, 100, 200, 400, 600, 800, 1000, 1500

[1] Deterministic 10 minutes.

[2] Exponential 10 minutes and seed equal to 0.

APPENDIX B

RESULTS OF SECTION 2

Table 67. Results from the comparison of the measures of dependence versus the MI calculated using the histogram-based method with fixed number of bins and probability density function for detecting the input with the greatest impact on the output, per number of bins and number of replications.

Bins	Number of replications	Distance correlation		Pearson correlation		R^2_{adj}	
		NIS	TIS	NIS	TIS	NIS	TIS
2	10	12.70%	12.70%	11.40%	15.20%	10.10%	11.40%
	20	25.30%	19.00%	22.80%	21.50%	19.00%	21.50%
	50	29.10%	29.10%	31.60%	31.60%	31.60%	31.60%
	100	35.40%	34.20%	34.20%	38.00%	34.20%	38.00%
	200	34.20%	43.00%	38.00%	44.30%	38.00%	44.30%
	400	40.50%	34.20%	35.40%	32.90%	35.40%	32.90%
	600	31.60%	29.10%	27.80%	27.80%	27.80%	27.80%
	800	43.00%	41.80%	41.80%	46.80%	41.80%	46.80%
	1000	41.80%	44.30%	44.30%	50.60%	44.30%	50.60%
	1500	40.50%	31.60%	39.20%	32.90%	39.20%	32.90%
2 Total		33.40%	31.90%	32.70%	34.20%	32.20%	33.80%
5	10	26.60%	36.70%	29.10%	38.00%	24.10%	35.40%
	20	36.70%	38.00%	36.70%	40.50%	32.90%	40.50%
	50	19.00%	26.60%	21.50%	27.80%	21.50%	27.80%
	100	40.50%	45.60%	41.80%	46.80%	41.80%	46.80%
	200	41.80%	40.50%	39.20%	39.20%	39.20%	39.20%
	400	43.00%	38.00%	38.00%	39.20%	38.00%	39.20%
	600	45.60%	38.00%	40.50%	45.60%	40.50%	45.60%
	800	36.70%	40.50%	38.00%	45.60%	38.00%	45.60%
	1000	32.90%	40.50%	34.20%	43.00%	34.20%	43.00%
	1500	38.00%	43.00%	43.00%	49.40%	43.00%	49.40%
5 Total		36.10%	38.70%	36.20%	41.50%	35.30%	41.30%
10	10	20.30%	19.00%	20.30%	16.50%	16.50%	12.70%
	20	35.40%	40.50%	41.80%	40.50%	39.20%	40.50%
	50	35.40%	46.80%	39.20%	46.80%	39.20%	46.80%
	100	59.50%	54.40%	54.40%	51.90%	54.40%	51.90%
	200	43.00%	34.20%	41.80%	43.00%	41.80%	43.00%
	400	44.30%	38.00%	34.20%	39.20%	34.20%	39.20%

	600	48.10%	43.00%	38.00%	46.80%	38.00%	46.80%
	800	43.00%	41.80%	34.20%	39.20%	34.20%	39.20%
	1000	40.50%	38.00%	30.40%	34.20%	30.40%	34.20%
	1500	38.00%	41.80%	39.20%	41.80%	39.20%	41.80%
	10 Total	40.80%	39.70%	37.30%	40.00%	36.70%	39.60%
25	10	16.50%	13.90%	12.70%	12.70%	12.70%	10.10%
	20	30.40%	31.60%	30.40%	32.90%	26.60%	32.90%
	50	45.60%	40.50%	43.00%	44.30%	43.00%	44.30%
	100	53.20%	45.60%	40.50%	43.00%	40.50%	43.00%
	200	49.40%	49.40%	53.20%	55.70%	53.20%	55.70%
	400	34.20%	36.70%	38.00%	39.20%	38.00%	39.20%
	600	38.00%	44.30%	43.00%	49.40%	43.00%	49.40%
	800	43.00%	44.30%	48.10%	50.60%	48.10%	50.60%
	1000	38.00%	41.80%	46.80%	48.10%	46.80%	48.10%
	1500	36.70%	38.00%	41.80%	45.60%	41.80%	45.60%
	25 Total	38.50%	38.60%	39.70%	42.20%	39.40%	41.90%
50	10	7.60%	6.30%	5.10%	3.80%	3.80%	2.50%
	20	29.10%	30.40%	31.60%	29.10%	26.60%	29.10%
	50	53.20%	43.00%	51.90%	44.30%	51.90%	44.30%
	100	46.80%	44.30%	44.30%	44.30%	44.30%	44.30%
	200	40.50%	34.20%	39.20%	35.40%	39.20%	35.40%
	400	34.20%	41.80%	51.90%	46.80%	51.90%	46.80%
	600	32.90%	44.30%	43.00%	41.80%	43.00%	41.80%
	800	48.10%	49.40%	51.90%	51.90%	51.90%	51.90%
	1000	48.10%	50.60%	51.90%	54.40%	51.90%	54.40%
	1500	51.90%	60.80%	54.40%	55.70%	54.40%	55.70%
	50 Total	39.20%	40.50%	42.50%	40.80%	41.90%	40.60%
100	10	11.40%	12.70%	8.90%	11.40%	7.60%	8.90%
	20	24.10%	26.60%	21.50%	26.60%	16.50%	26.60%
	50	39.20%	50.60%	39.20%	48.10%	39.20%	48.10%
	100	44.30%	32.90%	40.50%	31.60%	40.50%	31.60%
	200	38.00%	36.70%	40.50%	35.40%	40.50%	35.40%
	400	45.60%	38.00%	45.60%	36.70%	45.60%	36.70%
	600	44.30%	54.40%	46.80%	54.40%	46.80%	54.40%
	800	44.30%	49.40%	48.10%	50.60%	48.10%	50.60%
	1000	50.60%	58.20%	59.50%	57.00%	59.50%	57.00%
	1500	41.80%	50.60%	39.20%	44.30%	39.20%	44.30%
	100 Total	38.40%	41.00%	39.00%	39.60%	38.40%	39.40%
200	10	13.90%	13.90%	12.70%	13.90%	10.10%	10.10%

	20	17.70%	15.20%	16.50%	15.20%	11.40%	15.20%
	50	40.50%	44.30%	40.50%	45.60%	40.50%	45.60%
	100	50.60%	46.80%	50.60%	39.20%	50.60%	39.20%
	200	54.40%	41.80%	45.60%	46.80%	45.60%	46.80%
	400	41.80%	49.40%	50.60%	44.30%	50.60%	44.30%
	600	34.20%	53.20%	45.60%	48.10%	45.60%	48.10%
	800	51.90%	44.30%	50.60%	51.90%	50.60%	51.90%
	1000	51.90%	51.90%	55.70%	53.20%	55.70%	53.20%
	1500	46.80%	45.60%	44.30%	44.30%	44.30%	44.30%
	200 Total	40.40%	40.60%	41.30%	40.30%	40.50%	39.90%
500	10	12.70%	12.70%	11.40%	12.70%	8.90%	8.90%
	20	12.70%	13.90%	12.70%	13.90%	7.60%	13.90%
	50	40.50%	43.00%	36.70%	40.50%	36.70%	40.50%
	100	57.00%	46.80%	51.90%	38.00%	51.90%	38.00%
	200	45.60%	49.40%	43.00%	46.80%	43.00%	46.80%
	400	46.80%	51.90%	46.80%	49.40%	46.80%	49.40%
	600	38.00%	45.60%	44.30%	41.80%	44.30%	41.80%
	800	41.80%	46.80%	45.60%	35.40%	45.60%	35.40%
	1000	39.20%	39.20%	41.80%	34.20%	41.80%	34.20%
	1500	36.70%	36.70%	34.20%	32.90%	34.20%	32.90%
	500 Total	37.10%	38.60%	36.80%	34.60%	36.10%	34.20%
1000	10	12.70%	12.70%	11.40%	12.70%	8.90%	8.90%
	20	6.30%	6.30%	5.10%	6.30%	6.30%	6.30%
	50	38.00%	44.30%	36.70%	45.60%	36.70%	45.60%
	100	39.20%	34.20%	32.90%	31.60%	32.90%	31.60%
	200	41.80%	38.00%	40.50%	40.50%	40.50%	40.50%
	400	41.80%	48.10%	39.20%	51.90%	39.20%	51.90%
	600	43.00%	35.40%	40.50%	35.40%	40.50%	35.40%
	800	49.40%	39.20%	44.30%	35.40%	44.30%	35.40%
	1000	50.60%	39.20%	51.90%	32.90%	51.90%	32.90%
	1500	40.50%	43.00%	43.00%	39.20%	43.00%	39.20%
	1000 Total	36.30%	34.10%	34.60%	33.20%	34.40%	32.80%
2000	10	12.70%	12.70%	11.40%	12.70%	8.90%	8.90%
	20	12.70%	12.70%	12.70%	12.70%	7.60%	12.70%
	50	17.70%	22.80%	16.50%	25.30%	16.50%	25.30%
	100	50.60%	32.90%	35.40%	38.00%	35.40%	38.00%
	200	38.00%	36.70%	39.20%	39.20%	39.20%	39.20%
	400	48.10%	41.80%	38.00%	38.00%	38.00%	38.00%
	600	57.00%	43.00%	55.70%	43.00%	55.70%	43.00%

	800	53.20%	48.10%	54.40%	38.00%	54.40%	38.00%
	1000	45.60%	49.40%	40.50%	49.40%	40.50%	49.40%
	1500	45.60%	36.70%	43.00%	36.70%	43.00%	36.70%
	2000 Total	38.10%	33.70%	34.70%	33.30%	33.90%	32.90%

Table 68. Results from the comparison of the measures of dependence versus the MI calculated using the histogram-based method with fixed number of bins and probability density function for detecting the input with the least impact on the output, per number of bins and number of replications.

Bins	**Number of replications**	**Distance correlation**		**Pearson correlation**		R^2_{adj}	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
2	10	12.70%	15.20%	1.30%	8.90%	1.30%	2.50%
	20	32.90%	29.10%	21.50%	26.60%	20.30%	21.50%
	50	31.60%	35.40%	35.40%	36.70%	31.60%	36.70%
	100	39.20%	41.80%	38.00%	49.40%	35.40%	49.40%
	200	40.50%	44.30%	43.00%	45.60%	43.00%	45.60%
	400	46.80%	41.80%	40.50%	39.20%	40.50%	39.20%
	600	39.20%	35.40%	35.40%	35.40%	35.40%	35.40%
	800	49.40%	51.90%	48.10%	57.00%	48.10%	57.00%
	1000	46.80%	51.90%	48.10%	57.00%	48.10%	57.00%
	1500	39.20%	40.50%	36.70%	39.20%	36.70%	39.20%
	2 Total	37.80%	38.70%	34.80%	39.50%	34.10%	38.40%
5	10	36.70%	40.50%	25.30%	38.00%	25.30%	31.60%
	20	43.00%	45.60%	32.90%	44.30%	31.60%	39.20%
	50	27.80%	31.60%	27.80%	32.90%	24.10%	32.90%
	100	45.60%	51.90%	46.80%	55.70%	44.30%	55.70%
	200	49.40%	51.90%	46.80%	50.60%	46.80%	50.60%
	400	51.90%	46.80%	46.80%	46.80%	46.80%	46.80%
	600	51.90%	41.80%	48.10%	50.60%	48.10%	50.60%
	800	43.00%	43.00%	43.00%	49.40%	43.00%	49.40%
	1000	36.70%	43.00%	38.00%	48.10%	38.00%	48.10%
	1500	40.50%	44.30%	45.60%	51.90%	45.60%	51.90%
	5 Total	42.70%	44.10%	40.10%	46.80%	39.40%	45.70%
10	10	22.80%	25.30%	11.40%	15.20%	12.70%	8.90%
	20	44.30%	48.10%	39.20%	43.00%	38.00%	38.00%
	50	43.00%	50.60%	44.30%	48.10%	40.50%	48.10%
	100	59.50%	59.50%	55.70%	59.50%	53.20%	59.50%
	200	53.20%	46.80%	51.90%	57.00%	51.90%	57.00%

	400	49.40%	44.30%	39.20%	46.80%	39.20%	46.80%
	600	55.70%	48.10%	45.60%	51.90%	45.60%	51.90%
	800	48.10%	44.30%	39.20%	43.00%	39.20%	43.00%
	1000	46.80%	43.00%	36.70%	40.50%	36.70%	40.50%
	1500	44.30%	49.40%	45.60%	49.40%	45.60%	49.40%
10 Total		46.70%	45.90%	40.90%	45.40%	40.30%	44.30%
25	10	19.00%	17.70%	6.30%	11.40%	7.60%	5.10%
	20	40.50%	38.00%	29.10%	32.90%	27.80%	27.80%
	50	49.40%	40.50%	48.10%	45.60%	44.30%	45.60%
	100	55.70%	49.40%	43.00%	46.80%	40.50%	46.80%
	200	51.90%	48.10%	58.20%	55.70%	58.20%	55.70%
	400	34.20%	38.00%	38.00%	40.50%	38.00%	40.50%
	600	46.80%	50.60%	49.40%	53.20%	49.40%	53.20%
	800	49.40%	49.40%	54.40%	57.00%	54.40%	57.00%
	1000	45.60%	50.60%	54.40%	57.00%	54.40%	57.00%
	1500	44.30%	43.00%	50.60%	50.60%	50.60%	50.60%
25 Total		43.70%	42.50%	43.20%	45.10%	42.50%	43.90%
50	10	12.70%	13.90%	1.30%	7.60%	2.50%	1.30%
	20	34.20%	40.50%	25.30%	32.90%	24.10%	27.80%
	50	55.70%	49.40%	55.70%	50.60%	51.90%	50.60%
	100	49.40%	50.60%	46.80%	53.20%	44.30%	53.20%
	200	48.10%	43.00%	44.30%	43.00%	44.30%	43.00%
	400	39.20%	50.60%	58.20%	53.20%	58.20%	53.20%
	600	41.80%	53.20%	51.90%	49.40%	51.90%	49.40%
	800	57.00%	55.70%	62.00%	57.00%	62.00%	57.00%
	1000	57.00%	58.20%	62.00%	62.00%	62.00%	62.00%
	1500	53.20%	63.30%	57.00%	57.00%	57.00%	57.00%
50 Total		44.80%	47.80%	46.50%	46.60%	45.80%	45.40%
100	10	15.20%	15.20%	3.80%	8.90%	3.80%	2.50%
	20	22.80%	30.40%	11.40%	25.30%	10.10%	20.30%
	50	44.30%	57.00%	45.60%	54.40%	41.80%	54.40%
	100	46.80%	36.70%	45.60%	35.40%	43.00%	35.40%
	200	45.60%	43.00%	46.80%	41.80%	46.80%	41.80%
	400	46.80%	46.80%	48.10%	46.80%	48.10%	46.80%
	600	45.60%	59.50%	49.40%	59.50%	49.40%	59.50%
	800	53.20%	58.20%	58.20%	58.20%	58.20%	58.20%
	1000	53.20%	64.60%	64.60%	63.30%	64.60%	63.30%
	1500	50.60%	60.80%	50.60%	53.20%	50.60%	53.20%
100 Total		42.40%	47.20%	42.40%	44.70%	41.60%	43.50%

200	10	15.20%	15.20%	3.80%	8.90%	3.80%	2.50%
	20	17.70%	19.00%	7.60%	15.20%	6.30%	10.10%
	50	45.60%	53.20%	44.30%	51.90%	40.50%	51.90%
	100	54.40%	51.90%	54.40%	48.10%	51.90%	48.10%
	200	58.20%	49.40%	51.90%	53.20%	51.90%	53.20%
	400	49.40%	51.90%	57.00%	49.40%	57.00%	49.40%
	600	43.00%	55.70%	53.20%	50.60%	53.20%	50.60%
	800	58.20%	54.40%	58.20%	62.00%	58.20%	62.00%
	1000	53.20%	62.00%	58.20%	62.00%	58.20%	62.00%
	1500	51.90%	55.70%	50.60%	53.20%	50.60%	53.20%
200 Total		44.70%	46.80%	43.90%	45.40%	43.20%	44.30%
500	10	12.70%	12.70%	1.30%	6.30%	1.30%	0.00%
	20	12.70%	12.70%	3.80%	8.90%	2.50%	3.80%
	50	43.00%	44.30%	39.20%	44.30%	35.40%	44.30%
	100	58.20%	50.60%	54.40%	40.50%	51.90%	40.50%
	200	48.10%	58.20%	46.80%	54.40%	46.80%	54.40%
	400	51.90%	62.00%	54.40%	58.20%	54.40%	58.20%
	600	44.30%	55.70%	50.60%	53.20%	50.60%	53.20%
	800	46.80%	59.50%	49.40%	50.60%	49.40%	50.60%
	1000	48.10%	49.40%	49.40%	46.80%	49.40%	46.80%
	1500	45.60%	49.40%	44.30%	46.80%	44.30%	46.80%
500 Total		41.10%	45.40%	39.40%	41.00%	38.60%	39.90%
1000	10	12.70%	12.70%	1.30%	6.30%	1.30%	0.00%
	20	12.70%	13.90%	3.80%	10.10%	2.50%	5.10%
	50	43.00%	48.10%	41.80%	51.90%	38.00%	51.90%
	100	45.60%	43.00%	39.20%	44.30%	36.70%	44.30%
	200	44.30%	48.10%	41.80%	45.60%	41.80%	45.60%
	400	49.40%	55.70%	45.60%	59.50%	45.60%	59.50%
	600	50.60%	45.60%	45.60%	44.30%	45.60%	44.30%
	800	58.20%	46.80%	50.60%	44.30%	50.60%	44.30%
	1000	55.70%	48.10%	55.70%	43.00%	55.70%	43.00%
	1500	48.10%	54.40%	50.60%	49.40%	50.60%	49.40%
1000 Total		42.00%	41.60%	37.60%	39.90%	36.80%	38.70%
2000	10	12.70%	12.70%	1.30%	6.30%	1.30%	0.00%
	20	12.70%	12.70%	3.80%	8.90%	2.50%	3.80%
	50	19.00%	21.50%	19.00%	25.30%	15.20%	25.30%
	100	55.70%	39.20%	41.80%	44.30%	39.20%	44.30%
	200	41.80%	46.80%	43.00%	46.80%	43.00%	46.80%
	400	55.70%	49.40%	46.80%	44.30%	46.80%	44.30%

	600	59.50%	57.00%	54.40%	54.40%	54.40%	54.40%
	800	53.20%	57.00%	51.90%	46.80%	51.90%	46.80%
	1000	54.40%	53.20%	45.60%	51.90%	45.60%	51.90%
	1500	50.60%	44.30%	46.80%	43.00%	46.80%	43.00%
2000 Total		41.50%	39.40%	35.40%	37.20%	34.70%	36.10%

Table 69. Results from the comparison of the measures of dependence versus the MI calculated using the histogram-based method with optimum number of bins and probability density function for detecting the input with the greatest impact on the output, per number of bins and number of replications.

Bins rule	**Number of replications**	**Distance correlation**		**Pearson correlation**		R^2_{adj}	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
FD	10	32.90%	44.30%	44.30%	51.90%	36.70%	48.10%
	20	57.00%	59.50%	58.20%	60.80%	55.70%	60.80%
	50	44.30%	53.20%	48.10%	57.00%	48.10%	57.00%
	100	60.80%	55.70%	67.10%	62.00%	67.10%	62.00%
	200	45.60%	58.20%	73.40%	68.40%	73.40%	68.40%
	400	40.50%	51.90%	59.50%	59.50%	59.50%	59.50%
	600	46.80%	55.70%	69.60%	63.30%	69.60%	63.30%
	800	49.40%	60.80%	63.30%	64.60%	63.30%	64.60%
	1000	54.40%	62.00%	62.00%	65.80%	62.00%	65.80%
	1500	57.00%	64.60%	60.80%	55.70%	60.80%	55.70%
FD Total		48.90%	56.60%	60.60%	60.90%	59.60%	60.50%
Scott	10	60.80%	59.50%	70.90%	64.60%	69.60%	62.00%
	20	51.90%	60.80%	58.20%	64.60%	59.50%	64.60%
	50	69.60%	74.70%	75.90%	81.00%	75.90%	81.00%
	100	55.70%	64.60%	62.00%	65.80%	62.00%	65.80%
	200	54.40%	57.00%	63.30%	60.80%	63.30%	60.80%
	400	41.80%	59.50%	68.40%	67.10%	68.40%	67.10%
	600	41.80%	57.00%	59.50%	62.00%	59.50%	62.00%
	800	49.40%	63.30%	64.60%	70.90%	64.60%	70.90%
	1000	44.30%	62.00%	50.60%	62.00%	50.60%	62.00%
	1500	50.60%	57.00%	51.90%	55.70%	51.90%	55.70%
Scott Total		52.00%	61.50%	62.50%	65.40%	62.50%	65.20%
Sturges	10	6.30%	10.10%	6.30%	11.40%	5.10%	7.60%
	20	49.40%	49.40%	55.70%	50.60%	57.00%	50.60%
	50	57.00%	68.40%	63.30%	69.60%	63.30%	69.60%

	100	51.90%	58.20%	60.80%	62.00%	60.80%	62.00%
	200	44.30%	55.70%	59.50%	64.60%	59.50%	64.60%
	400	48.10%	65.80%	58.20%	69.60%	58.20%	69.60%
	600	49.40%	55.70%	43.00%	64.60%	43.00%	64.60%
	800	55.70%	51.90%	36.70%	45.60%	36.70%	45.60%
	1000	51.90%	53.20%	34.20%	48.10%	34.20%	48.10%
	1500	48.10%	51.90%	41.80%	53.20%	41.80%	53.20%
Sturges Total		46.20%	52.00%	45.90%	53.90%	45.90%	53.50%

Table 70. Results from the comparison of the measures of dependence versus the MI calculated using the histogram-based method with optimum number of bins and probability density function for detecting the input with the least impact on the output, per number of bins and number of replications.

Bins rule	Number of replications	Distance correlation		Pearson correlation		R^2_{adj}	
		NIS	TIS	NIS	TIS	NIS	TIS
FD	10	49.40%	48.10%	41.80%	45.60%	40.50%	39.20%
	20	67.10%	72.20%	62.00%	70.90%	60.80%	65.80%
	50	55.70%	62.00%	57.00%	64.60%	54.40%	65.80%
	100	63.30%	62.00%	70.90%	68.40%	64.60%	72.20%
	200	53.20%	67.10%	82.30%	78.50%	81.00%	84.80%
	400	46.80%	59.50%	67.10%	67.10%	65.80%	73.40%
	600	48.10%	57.00%	72.20%	64.60%	70.90%	69.60%
	800	55.70%	65.80%	69.60%	69.60%	67.10%	72.20%
	1000	54.40%	64.60%	64.60%	67.10%	65.80%	65.80%
	1500	59.50%	64.60%	64.60%	57.00%	65.80%	58.20%
FD Total		55.30%	62.30%	65.20%	65.30%	63.70%	66.70%
Scott	10	62.00%	62.00%	57.00%	58.20%	57.00%	51.90%
	20	63.30%	68.40%	62.00%	69.60%	59.50%	65.80%
	50	72.20%	75.90%	78.50%	83.50%	72.20%	84.80%
	100	58.20%	64.60%	67.10%	68.40%	62.00%	70.90%
	200	55.70%	58.20%	67.10%	64.60%	67.10%	64.60%
	400	45.60%	62.00%	73.40%	69.60%	72.20%	72.20%
	600	51.90%	64.60%	67.10%	67.10%	64.60%	67.10%
	800	50.60%	64.60%	68.40%	73.40%	60.80%	74.70%
	1000	45.60%	62.00%	55.70%	62.00%	54.40%	65.80%
	1500	51.90%	58.20%	53.20%	58.20%	49.40%	55.70%
Scott Total		55.70%	64.10%	64.90%	67.50%	61.90%	67.30%

	10	13.90%	15.20%	2.50%	8.90%	2.50%	2.50%
	20	60.80%	65.80%	60.80%	62.00%	67.10%	63.30%
	50	64.60%	73.40%	70.90%	75.90%	68.40%	75.90%
	100	59.50%	67.10%	70.90%	73.40%	68.40%	77.20%
Sturges	200	53.20%	60.80%	69.60%	72.20%	70.90%	72.20%
	400	53.20%	70.90%	65.80%	75.90%	65.80%	74.70%
	600	60.80%	65.80%	57.00%	75.90%	57.00%	74.70%
	800	65.80%	59.50%	45.60%	53.20%	44.30%	55.70%
	1000	62.00%	58.20%	45.60%	55.70%	43.00%	57.00%
	1500	58.20%	57.00%	50.60%	59.50%	49.40%	59.50%
Sturges Total		55.20%	59.40%	53.90%	61.30%	53.70%	61.30%

Table 71. Results from the comparison of the distance correlation versus the MI calculated using the histogram-based method with fixed number of bins and discrete empirical distribution for detecting the input with the greatest impact on the output, per number of bins, number of replications, and normalization method.

Bins	**Number of replications**	**MI [%]**		NMI_{arith} **[%]**		NMI_{geom} **[%]**		NMI_{joint} **[%]**		NMI_{theor} **[%]**	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
	10	1.3	10.1	1.3	10.1	1.3	10.1	1.3	10.1	1.3	10.1
	20	67.1	68.4	67.1	68.4	67.1	68.4	67.1	68.4	67.1	68.4
	50	60.8	65.8	60.8	65.8	60.8	65.8	60.8	65.8	60.8	65.8
	100	50.6	65.8	53.2	64.6	53.2	64.6	53.2	64.6	50.6	65.8
2	200	46.8	64.6	44.3	64.6	44.3	64.6	44.3	64.6	46.8	64.6
	400	44.3	60.8	44.3	60.8	44.3	60.8	44.3	60.8	44.3	60.8
	600	43.0	62.0	44.3	62.0	44.3	62.0	44.3	62.0	43.0	62.0
	800	46.8	59.5	46.8	59.5	46.8	59.5	46.8	59.5	46.8	59.5
	1000	49.4	60.8	49.4	60.8	49.4	60.8	49.4	60.8	49.4	60.8
	1500	50.6	63.3	50.6	63.3	50.6	63.3	50.6	63.3	50.6	63.3
2 Total		46.1	58.1	46.2	58.0	46.2	58.0	46.2	58.0	46.1	58.1
	10	43.0	53.2	46.8	57.0	46.8	57.0	46.8	57.0	40.5	51.9
	20	64.6	65.8	63.3	65.8	63.3	65.8	63.3	65.8	64.6	65.8
	50	63.3	64.6	60.8	62.0	60.8	62.0	60.8	62.0	63.3	64.6
	100	54.4	69.6	55.7	70.9	55.7	70.9	55.7	70.9	54.4	69.6
5	200	51.9	70.9	51.9	72.2	51.9	72.2	51.9	72.2	51.9	70.9
	400	43.0	63.3	44.3	62.0	44.3	62.0	44.3	62.0	43.0	63.3
	600	43.0	62.0	43.0	60.8	43.0	60.8	43.0	60.8	43.0	62.0
	800	44.3	59.5	45.6	57.0	45.6	58.2	45.6	57.0	44.3	59.5
	1000	46.8	59.5	48.1	59.5	48.1	59.5	48.1	59.5	46.8	59.5

	1500	44.3	60.8	46.8	60.8	46.8	62.0	46.8	60.8	44.3	60.8
5 Total		49.9	62.9	50.6	62.8	50.6	63.0	50.6	62.8	49.6	62.8
10	10	31.6	49.4	31.6	50.6	31.6	50.6	31.6	50.6	31.6	49.4
	20	58.2	72.2	60.8	70.9	60.8	70.9	60.8	70.9	58.2	72.2
	50	53.2	65.8	51.9	67.1	51.9	67.1	51.9	67.1	53.2	65.8
	100	59.5	64.6	54.4	60.8	54.4	60.8	54.4	60.8	59.5	64.6
	200	44.3	55.7	44.3	53.2	44.3	51.9	44.3	53.2	44.3	55.7
	400	45.6	65.8	45.6	64.6	45.6	64.6	45.6	64.6	45.6	65.8
	600	44.3	63.3	40.5	60.8	40.5	60.8	40.5	60.8	44.3	63.3
	800	51.9	62.0	54.4	63.3	55.7	63.3	54.4	63.3	51.9	62.0
	1000	54.4	62.0	54.4	60.8	54.4	60.8	54.4	60.8	54.4	62.0
	1500	53.2	60.8	53.2	62.0	53.2	62.0	53.2	62.0	53.2	60.8
10 Total		49.6	62.2	49.1	61.4	49.2	61.3	49.1	61.4	49.6	62.2
25	10	15.2	17.7	16.5	19.0	16.5	19.0	15.2	17.7	15.2	17.7
	20	63.3	64.6	59.5	64.6	59.5	64.6	59.5	64.6	63.3	64.6
	50	57.0	68.4	57.0	70.9	57.0	70.9	57.0	70.9	57.0	68.4
	100	51.9	64.6	49.4	59.5	49.4	59.5	49.4	59.5	51.9	64.6
	200	48.1	58.2	48.1	57.0	49.4	57.0	48.1	57.0	48.1	58.2
	400	57.0	57.0	57.0	60.8	57.0	60.8	57.0	60.8	57.0	57.0
	600	54.4	57.0	54.4	57.0	54.4	57.0	54.4	57.0	54.4	57.0
	800	40.5	59.5	44.3	59.5	44.3	59.5	44.3	59.5	40.5	59.5
	1000	43.0	65.8	49.4	64.6	49.4	64.6	49.4	64.6	43.0	65.8
	1500	50.6	63.3	54.4	62.0	54.4	62.0	54.4	62.0	50.6	63.3
25 Total		48.1	57.6	49.0	57.5	49.1	57.5	48.9	57.3	48.1	57.6
50	10	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9
	20	36.7	40.5	39.2	58.2	39.2	58.2	39.2	58.2	36.7	40.5
	50	51.9	55.7	51.9	64.6	51.9	64.6	51.9	64.6	51.9	55.7
	100	53.2	63.3	51.9	60.8	51.9	60.8	51.9	60.8	53.2	63.3
	200	51.9	55.7	53.2	55.7	53.2	55.7	53.2	55.7	51.9	55.7
	400	51.9	48.1	53.2	51.9	53.2	51.9	53.2	51.9	51.9	48.1
	600	51.9	57.0	54.4	54.4	54.4	54.4	54.4	54.4	51.9	57.0
	800	36.7	54.4	43.0	50.6	41.8	50.6	43.0	50.6	36.7	54.4
	1000	39.2	59.5	44.3	58.2	44.3	58.2	44.3	58.2	39.2	59.5
	1500	46.8	55.7	49.4	53.2	49.4	53.2	49.4	53.2	46.8	55.7
50 Total		42.9	49.9	44.9	51.6	44.8	51.6	44.9	51.6	42.9	49.9
100	10	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7
	20	35.4	35.4	39.2	39.2	39.2	39.2	39.2	39.2	35.4	35.4
	50	49.4	53.2	43.0	57.0	43.0	57.0	43.0	57.0	49.4	53.2
	100	50.6	62.0	44.3	55.7	44.3	55.7	44.3	55.7	50.6	62.0

	200	43.0	59.5	44.3	59.5	44.3	59.5	44.3	59.5	43.0	59.5
	400	44.3	49.4	50.6	48.1	51.9	48.1	50.6	48.1	44.3	49.4
	600	40.5	53.2	50.6	50.6	50.6	49.4	50.6	50.6	40.5	53.2
	800	48.1	54.4	48.1	51.9	48.1	51.9	48.1	51.9	48.1	54.4
	1000	55.7	53.2	55.7	54.4	55.7	53.2	55.7	54.4	55.7	53.2
	1500	48.1	51.9	45.6	48.1	45.6	48.1	45.6	48.1	48.1	51.9
100 Total		42.7	48.5	43.3	47.7	43.4	47.5	43.3	47.7	42.7	48.5
	10	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9
	20	16.5	16.5	16.5	16.5	16.5	16.5	16.5	16.5	16.5	16.5
	50	54.4	46.8	51.9	51.9	51.9	51.9	51.9	53.2	54.4	46.8
	100	53.2	60.8	46.8	63.3	45.6	63.3	46.8	63.3	53.2	60.8
200	200	50.6	57.0	50.6	48.1	50.6	48.1	50.6	48.1	50.6	57.0
	400	49.4	55.7	67.1	51.9	67.1	51.9	67.1	51.9	49.4	55.7
	600	45.6	54.4	54.4	62.0	54.4	62.0	54.4	62.0	45.6	54.4
	800	54.4	53.2	51.9	58.2	51.9	58.2	51.9	58.2	54.4	53.2
	1000	54.4	57.0	46.8	45.6	46.8	45.6	46.8	45.6	54.4	57.0
	1500	57.0	57.0	48.1	53.2	48.1	53.2	48.1	53.2	57.0	57.0
200 Total		44.9	47.2	44.8	46.5	44.7	46.5	44.8	46.6	44.9	47.2
	10	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	20	12.7	13.9	12.7	13.9	12.7	13.9	12.7	13.9	12.7	13.9
	50	40.5	51.9	49.4	60.8	49.4	60.8	49.4	60.8	40.5	51.9
	100	62.0	63.3	54.4	67.1	54.4	67.1	54.4	67.1	62.0	63.3
500	200	45.6	64.6	48.1	62.0	48.1	62.0	48.1	62.0	45.6	64.6
	400	43.0	57.0	45.6	58.2	45.6	58.2	45.6	58.2	43.0	57.0
	600	34.2	58.2	43.0	60.8	43.0	60.8	43.0	60.8	34.2	58.2
	800	40.5	63.3	41.8	55.7	41.8	55.7	41.8	55.7	40.5	63.3
	1000	46.8	58.2	39.2	49.4	39.2	48.1	39.2	49.4	46.8	58.2
	1500	51.9	62.0	48.1	59.5	48.1	59.5	48.1	59.5	51.9	62.0
500 Total		39.0	50.5	39.5	50.0	39.5	49.9	39.5	50.0	39.0	50.5
	10	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	20	12.7	13.9	12.7	13.9	12.7	13.9	12.7	13.9	12.7	13.9
	50	39.2	43.0	44.3	50.6	44.3	50.6	44.3	50.6	39.2	43.0
	100	54.4	60.8	54.4	59.5	54.4	59.5	54.4	59.5	54.4	60.8
1000	200	59.5	62.0	54.4	60.8	54.4	60.8	54.4	60.8	59.5	62.0
	400	45.6	51.9	40.5	48.1	40.5	48.1	40.5	48.1	45.6	51.9
	600	40.5	55.7	44.3	57.0	44.3	57.0	44.3	57.0	40.5	55.7
	800	44.3	55.7	44.3	51.9	44.3	51.9	44.3	51.9	44.3	55.7
	1000	49.4	57.0	38.0	54.4	38.0	54.4	38.0	54.4	49.4	57.0
	1500	48.1	59.5	50.6	63.3	50.6	63.3	50.6	63.3	48.1	59.5

1000 Total		40.6	47.2	39.6	47.2	39.6	47.2	39.6	47.2	40.6	47.2
2000	10	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	20	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7
	50	30.4	32.9	31.6	32.9	31.6	32.9	31.6	32.9	30.4	32.9
	100	54.4	53.2	53.2	51.9	53.2	51.9	53.2	51.9	53.2	53.2
	200	58.2	59.5	57.0	57.0	57.0	57.0	57.0	57.0	58.2	59.5
	400	45.6	55.7	44.3	50.6	44.3	50.6	44.3	50.6	45.6	55.7
	600	40.5	54.4	39.2	55.7	39.2	55.7	39.2	55.7	40.5	54.4
	800	40.5	58.2	48.1	58.2	48.1	57.0	48.1	58.2	40.5	58.2
	1000	44.3	60.8	45.6	55.7	45.6	55.7	45.6	55.7	44.3	60.8
	1500	51.9	55.7	45.6	54.4	45.6	54.4	45.6	54.4	51.9	55.7
2000 Total		39.0	45.6	38.9	44.2	38.9	44.1	38.9	44.2	38.9	45.6

Table 72. Results from the comparison of the distance correlation versus the MI calculated using the histogram-based method with fixed number of bins and discrete empirical distribution for detecting the input with the least impact on the output, per number of bins, number of replications, and normalization method.

Bins	**Number of replications**	**MI [%]**		NMI_{arith} **[%]**		NMI_{geom} **[%]**		NMI_{joint} **[%]**		NMI_{theor} **[%]**	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
2	10	15.2	13.9	15.2	13.9	15.2	13.9	15.2	13.9	15.2	13.9
	20	72.2	74.7	72.2	74.7	72.2	74.7	72.2	74.7	72.2	74.7
	50	63.3	67.1	63.3	67.1	63.3	67.1	63.3	67.1	63.3	67.1
	100	59.5	74.7	62.0	73.4	62.0	73.4	62.0	73.4	59.5	74.7
	200	57.0	70.9	55.7	70.9	55.7	70.9	55.7	70.9	57.0	70.9
	400	50.6	63.3	49.4	63.3	49.4	63.3	49.4	63.3	50.6	63.3
	600	51.9	64.6	51.9	64.6	51.9	64.6	51.9	64.6	51.9	64.6
	800	57.0	65.8	57.0	65.8	57.0	65.8	57.0	65.8	57.0	65.8
	1000	55.7	65.8	55.7	65.8	55.7	65.8	55.7	65.8	55.7	65.8
	1500	57.0	65.8	57.0	65.8	57.0	65.8	57.0	65.8	57.0	65.8
2 Total		53.9	62.7	53.9	62.5	53.9	62.5	53.9	62.5	53.9	62.7
5	10	53.2	62.0	55.7	63.3	55.7	63.3	55.7	63.3	50.6	59.5
	20	70.9	74.7	70.9	74.7	70.9	74.7	70.9	74.7	70.9	74.7
	50	65.8	65.8	64.6	65.8	64.6	65.8	64.6	65.8	65.8	65.8
	100	63.3	74.7	64.6	75.9	64.6	75.9	64.6	75.9	63.3	74.7
	200	57.0	75.9	59.5	77.2	59.5	77.2	59.5	77.2	57.0	75.9
	400	49.4	67.1	50.6	65.8	50.6	65.8	50.6	65.8	49.4	67.1
	600	53.2	65.8	53.2	64.6	53.2	64.6	53.2	64.6	53.2	65.8
	800	53.2	67.1	54.4	62.0	54.4	63.3	54.4	62.0	53.2	67.1

	1000	51.9	67.1	50.6	64.6	50.6	64.6	50.6	64.6	51.9	67.1
	1500	51.9	64.6	53.2	63.3	53.2	64.6	53.2	63.3	51.9	64.6
5 Total		57.0	68.5	57.7	67.7	57.7	68.0	57.7	67.7	56.7	68.2
10	10	41.8	51.9	41.8	53.2	41.8	53.2	41.8	53.2	41.8	51.9
	20	63.3	70.9	64.6	70.9	64.6	70.9	64.6	70.9	63.3	70.9
	50	59.5	68.4	60.8	70.9	60.8	70.9	60.8	70.9	59.5	68.4
	100	64.6	70.9	58.2	67.1	58.2	67.1	58.2	67.1	64.6	70.9
	200	50.6	60.8	54.4	60.8	54.4	60.8	54.4	60.8	50.6	60.8
	400	51.9	69.6	54.4	69.6	54.4	69.6	54.4	69.6	51.9	69.6
	600	53.2	64.6	51.9	64.6	51.9	64.6	51.9	64.6	53.2	64.6
	800	60.8	65.8	64.6	68.4	65.8	68.4	64.6	68.4	60.8	65.8
	1000	64.6	67.1	64.6	64.6	64.6	64.6	64.6	64.6	64.6	67.1
	1500	65.8	64.6	65.8	64.6	65.8	64.6	65.8	64.6	65.8	64.6
10 Total		57.6	65.4	58.1	65.4	58.2	65.4	58.1	65.4	57.6	65.4
25	10	17.7	20.3	17.7	20.3	17.7	20.3	17.7	20.3	17.7	20.3
	20	69.6	67.1	67.1	70.9	67.1	70.9	67.1	70.9	69.6	67.1
	50	65.8	69.6	65.8	70.9	65.8	70.9	65.8	70.9	65.8	69.6
	100	58.2	64.6	53.2	63.3	53.2	63.3	53.2	63.3	58.2	64.6
	200	57.0	58.2	55.7	59.5	57.0	59.5	55.7	59.5	57.0	58.2
	400	63.3	59.5	65.8	64.6	65.8	64.6	65.8	64.6	63.3	59.5
	600	63.3	63.3	65.8	63.3	65.8	63.3	65.8	63.3	63.3	63.3
	800	53.2	68.4	57.0	65.8	57.0	65.8	57.0	65.8	53.2	68.4
	1000	53.2	68.4	58.2	67.1	58.2	67.1	58.2	67.1	53.2	68.4
	1500	55.7	64.6	58.2	63.3	58.2	63.3	58.2	63.3	55.7	64.6
25 Total		55.7	60.4	56.5	60.9	56.6	60.9	56.5	60.9	55.7	60.4
50	10	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	20	43.0	46.8	46.8	65.8	46.8	65.8	46.8	65.8	43.0	46.8
	50	60.8	64.6	57.0	73.4	57.0	73.4	57.0	73.4	60.8	64.6
	100	55.7	63.3	57.0	58.2	57.0	57.0	57.0	58.2	55.7	63.3
	200	53.2	58.2	57.0	57.0	57.0	57.0	57.0	57.0	53.2	58.2
	400	53.2	50.6	54.4	53.2	54.4	53.2	54.4	53.2	53.2	50.6
	600	53.2	55.7	55.7	57.0	55.7	57.0	55.7	57.0	53.2	55.7
	800	49.4	60.8	54.4	58.2	53.2	58.2	54.4	58.2	49.4	60.8
	1000	46.8	62.0	53.2	60.8	53.2	60.8	53.2	60.8	46.8	62.0
	1500	50.6	60.8	55.7	58.2	55.7	58.2	55.7	58.2	50.6	60.8
50 Total		47.8	53.5	50.4	55.4	50.3	55.3	50.4	55.4	47.8	53.5
100	10	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2
	20	38.0	38.0	41.8	41.8	41.8	41.8	41.8	41.8	38.0	38.0
	50	54.4	55.7	48.1	58.2	48.1	58.2	48.1	58.2	54.4	55.7

	100	51.9	60.8	45.6	55.7	45.6	55.7	45.6	55.7	51.9	60.8
	200	49.4	62.0	51.9	64.6	51.9	64.6	51.9	64.6	49.4	62.0
	400	51.9	54.4	57.0	53.2	57.0	53.2	57.0	53.2	51.9	54.4
	600	46.8	60.8	57.0	58.2	57.0	58.2	57.0	58.2	46.8	60.8
	800	57.0	58.2	60.8	58.2	60.8	58.2	60.8	58.2	57.0	58.2
	1000	57.0	55.7	59.5	59.5	59.5	58.2	59.5	59.5	57.0	55.7
	1500	55.7	54.4	57.0	57.0	57.0	57.0	57.0	57.0	55.7	54.4
100 Total		47.7	51.5	49.4	52.2	49.4	52.0	49.4	52.2	47.7	51.5
	10	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2	15.2
	20	20.3	19.0	20.3	19.0	20.3	19.0	20.3	19.0	20.3	19.0
	50	51.9	50.6	51.9	55.7	51.9	57.0	51.9	55.7	51.9	50.6
	100	55.7	62.0	50.6	65.8	49.4	65.8	50.6	65.8	55.7	62.0
	200	53.2	57.0	54.4	53.2	54.4	53.2	54.4	53.2	53.2	57.0
200	400	53.2	58.2	68.4	54.4	68.4	54.4	68.4	54.4	53.2	58.2
	600	51.9	58.2	60.8	65.8	60.8	65.8	60.8	65.8	51.9	58.2
	800	60.8	59.5	59.5	62.0	59.5	62.0	59.5	62.0	60.8	59.5
	1000	53.2	59.5	54.4	49.4	54.4	49.4	54.4	49.4	53.2	59.5
	1500	64.6	59.5	58.2	55.7	58.2	55.7	58.2	55.7	64.6	59.5
200 Total		48.0	49.9	49.4	49.6	49.2	49.7	49.4	49.6	48.0	49.9
	10	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	20	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	50	41.8	57.0	51.9	63.3	51.9	63.3	51.9	63.3	41.8	57.0
	100	62.0	63.3	57.0	65.8	57.0	65.8	57.0	65.8	62.0	63.3
	200	49.4	67.1	51.9	67.1	51.9	67.1	51.9	67.1	49.4	67.1
500	400	51.9	57.0	50.6	58.2	50.6	58.2	50.6	58.2	51.9	57.0
	600	41.8	60.8	50.6	62.0	50.6	62.0	50.6	62.0	41.8	60.8
	800	49.4	68.4	48.1	65.8	48.1	65.8	48.1	65.8	49.4	68.4
	1000	55.7	65.8	45.6	58.2	45.6	57.0	45.6	58.2	55.7	65.8
	1500	55.7	64.6	54.4	65.8	54.4	65.8	54.4	65.8	55.7	64.6
500 Total		43.3	52.9	43.5	53.2	43.5	53.0	43.5	53.2	43.3	52.9
	10	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	20	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	50	41.8	53.2	46.8	57.0	46.8	57.0	46.8	57.0	41.8	53.2
	100	55.7	60.8	54.4	57.0	54.4	57.0	54.4	57.0	55.7	60.8
1000	200	59.5	63.3	55.7	59.5	55.7	59.5	55.7	59.5	59.5	63.3
	400	51.9	55.7	49.4	49.4	49.4	49.4	49.4	49.4	51.9	55.7
	600	46.8	60.8	50.6	60.8	50.6	60.8	50.6	60.8	46.8	60.8
	800	51.9	64.6	54.4	55.7	54.4	55.7	54.4	55.7	51.9	64.6
	1000	53.2	60.8	43.0	59.5	43.0	59.5	43.0	59.5	53.2	60.8

	1500	50.6	64.6	55.7	63.3	55.7	63.3	55.7	63.3	50.6	64.6
1000 Total		43.7	50.9	43.5	48.7	43.5	48.7	43.5	48.7	43.7	50.9
2000	10	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	20	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	50	29.1	31.6	30.4	30.4	30.4	30.4	30.4	30.4	29.1	31.6
	100	54.4	53.2	53.2	51.9	53.2	51.9	53.2	51.9	53.2	53.2
	200	62.0	65.8	63.3	64.6	63.3	64.6	63.3	64.6	62.0	67.1
	400	48.1	58.2	46.8	51.9	46.8	51.9	46.8	51.9	48.1	58.2
	600	48.1	63.3	45.6	59.5	45.6	59.5	45.6	59.5	48.1	63.3
	800	50.6	64.6	54.4	58.2	54.4	57.0	54.4	58.2	50.6	64.6
	1000	50.6	63.3	54.4	58.2	54.4	58.2	54.4	58.2	50.6	63.3
	1500	55.7	60.8	50.6	59.5	50.6	59.5	50.6	59.5	55.7	60.8
2000 Total		42.4	48.6	42.4	45.9	42.4	45.8	42.4	45.9	42.3	48.7

Table 73. Results from the comparison of the Pearson correlation versus the MI calculated using the histogram-based method with fixed number of bins and discrete empirical distribution for detecting the input with the greatest impact on the output, per number of bins, number of replications, and normalization method.

Bins	**Number of replications**	**MI [%]**		NMI_{arith} **[%]**		NMI_{geom} **[%]**		NMI_{joint} **[%]**		NMI_{theor} **[%]**	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
2	10	8.9	13.9	8.9	13.9	8.9	13.9	8.9	13.9	8.9	13.9
	20	67.1	67.1	67.1	67.1	67.1	67.1	67.1	67.1	67.1	67.1
	50	68.4	70.9	68.4	70.9	68.4	70.9	68.4	70.9	68.4	70.9
	100	75.9	82.3	78.5	81.0	78.5	81.0	78.5	81.0	75.9	82.3
	200	72.2	79.7	69.6	79.7	69.6	79.7	69.6	79.7	72.2	79.7
	400	58.2	75.9	58.2	75.9	58.2	75.9	58.2	75.9	58.2	75.9
	600	58.2	74.7	59.5	74.7	59.5	74.7	59.5	74.7	58.2	74.7
	800	63.3	70.9	63.3	70.9	63.3	70.9	63.3	70.9	63.3	70.9
	1000	67.1	70.9	67.1	70.9	67.1	70.9	67.1	70.9	67.1	70.9
	1500	68.4	74.7	68.4	74.7	68.4	74.7	68.4	74.7	68.4	74.7
2 Total		60.8	68.1	60.9	68.0	60.9	68.0	60.9	68.0	60.8	68.1
5	10	46.8	57.0	53.2	60.8	53.2	60.8	53.2	60.8	45.6	57.0
	20	73.4	69.6	72.2	69.6	72.2	69.6	72.2	69.6	73.4	69.6
	50	69.6	67.1	69.6	64.6	69.6	64.6	69.6	64.6	69.6	67.1
	100	72.2	77.2	70.9	75.9	70.9	75.9	70.9	75.9	72.2	77.2
	200	69.6	75.9	67.1	74.7	67.1	74.7	67.1	74.7	69.6	75.9
	400	57.0	73.4	55.7	72.2	55.7	72.2	55.7	72.2	57.0	73.4
	600	55.7	72.2	55.7	70.9	55.7	70.9	55.7	70.9	55.7	72.2

	800	63.3	70.9	64.6	70.9	64.6	69.6	64.6	70.9	63.3	70.9
	1000	64.6	69.6	65.8	72.2	65.8	72.2	65.8	72.2	64.6	69.6
	1500	58.2	67.1	60.8	67.1	60.8	65.8	60.8	67.1	58.2	67.1
5 Total		63.0	70.0	63.5	69.9	63.5	69.6	63.5	69.9	62.9	70.0
10	10	36.7	50.6	36.7	51.9	36.7	51.9	36.7	51.9	36.7	50.6
	20	62.0	74.7	64.6	74.7	64.6	74.7	64.6	74.7	62.0	74.7
	50	57.0	67.1	55.7	68.4	55.7	68.4	55.7	68.4	57.0	67.1
	100	55.7	65.8	55.7	62.0	55.7	62.0	55.7	62.0	55.7	65.8
	200	46.8	64.6	44.3	59.5	44.3	58.2	44.3	59.5	46.8	64.6
	400	44.3	63.3	44.3	62.0	44.3	62.0	44.3	62.0	44.3	63.3
	600	44.3	63.3	40.5	60.8	40.5	60.8	40.5	60.8	44.3	63.3
	800	45.6	58.2	43.0	59.5	44.3	59.5	43.0	59.5	45.6	58.2
	1000	46.8	62.0	44.3	58.2	44.3	58.2	44.3	58.2	46.8	62.0
	1500	44.3	64.6	41.8	63.3	41.8	63.3	41.8	63.3	44.3	64.6
10 Total		48.4	63.4	47.1	62.0	47.2	61.9	47.1	62.0	48.4	63.4
25	10	11.4	17.7	11.4	16.5	11.4	16.5	11.4	16.5	11.4	17.7
	20	62.0	67.1	58.2	67.1	58.2	67.1	58.2	67.1	62.0	67.1
	50	55.7	74.7	55.7	77.2	55.7	77.2	55.7	77.2	55.7	74.7
	100	46.8	57.0	44.3	54.4	44.3	54.4	44.3	54.4	46.8	57.0
	200	55.7	62.0	53.2	60.8	51.9	60.8	53.2	60.8	55.7	62.0
	400	49.4	58.2	46.8	54.4	46.8	54.4	46.8	54.4	49.4	58.2
	600	48.1	50.6	45.6	50.6	45.6	50.6	45.6	50.6	48.1	50.6
	800	45.6	57.0	44.3	54.4	44.3	54.4	44.3	54.4	45.6	57.0
	1000	44.3	65.8	45.6	62.0	45.6	62.0	45.6	62.0	44.3	65.8
	1500	51.9	60.8	50.6	59.5	50.6	59.5	50.6	59.5	51.9	60.8
25 Total		47.1	57.1	45.6	55.7	45.4	55.7	45.6	55.7	47.1	57.1
50	10	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3
	20	34.2	38.0	39.2	55.7	39.2	55.7	39.2	55.7	34.2	38.0
	50	48.1	54.4	48.1	63.3	48.1	63.3	48.1	63.3	48.1	54.4
	100	62.0	62.0	53.2	59.5	53.2	59.5	53.2	59.5	62.0	62.0
	200	64.6	65.8	55.7	65.8	55.7	65.8	55.7	65.8	64.6	65.8
	400	58.2	58.2	57.0	59.5	57.0	59.5	57.0	59.5	58.2	58.2
	600	59.5	59.5	57.0	57.0	57.0	57.0	57.0	57.0	59.5	59.5
	800	40.5	53.2	44.3	51.9	43.0	51.9	44.3	51.9	40.5	53.2
	1000	45.6	60.8	45.6	59.5	45.6	59.5	45.6	59.5	45.6	60.8
	1500	44.3	55.7	46.8	55.7	46.8	55.7	46.8	55.7	44.3	55.7
50 Total		45.8	51.4	44.8	53.4	44.7	53.4	44.8	53.4	45.8	51.4
100	10	8.9	11.4	8.9	11.4	8.9	11.4	8.9	11.4	8.9	11.4
	20	36.7	36.7	40.5	40.5	40.5	40.5	40.5	40.5	36.7	36.7

	50	49.4	51.9	46.8	55.7	46.8	55.7	46.8	55.7	49.4	51.9
	100	64.6	65.8	51.9	59.5	51.9	60.8	51.9	59.5	64.6	65.8
	200	59.5	57.0	55.7	59.5	55.7	59.5	55.7	59.5	59.5	57.0
	400	53.2	50.6	49.4	46.8	50.6	46.8	49.4	46.8	53.2	50.6
	600	55.7	49.4	50.6	46.8	50.6	45.6	50.6	46.8	55.7	49.4
	800	53.2	53.2	51.9	50.6	51.9	50.6	51.9	50.6	53.2	53.2
	1000	62.0	51.9	58.2	50.6	58.2	49.4	58.2	50.6	62.0	51.9
	1500	59.5	51.9	57.0	45.6	57.0	45.6	57.0	45.6	59.5	51.9
100 Total		50.3	48.0	47.1	46.7	47.2	46.6	47.1	46.7	50.3	48.0
	10	12.7	13.9	12.7	13.9	12.7	13.9	12.7	13.9	12.7	13.9
	20	16.5	16.5	16.5	16.5	16.5	16.5	16.5	16.5	16.5	16.5
	50	57.0	45.6	55.7	48.1	55.7	48.1	55.7	49.4	57.0	45.6
	100	57.0	54.4	46.8	55.7	48.1	55.7	46.8	55.7	57.0	54.4
200	200	55.7	58.2	53.2	51.9	53.2	51.9	53.2	51.9	55.7	58.2
	400	55.7	55.7	53.2	51.9	53.2	51.9	53.2	51.9	55.7	55.7
	600	53.2	54.4	49.4	57.0	49.4	57.0	49.4	57.0	53.2	54.4
	800	60.8	51.9	48.1	51.9	48.1	51.9	48.1	51.9	60.8	51.9
	1000	64.6	57.0	51.9	50.6	51.9	50.6	51.9	50.6	64.6	57.0
	1500	63.3	53.2	54.4	51.9	54.4	51.9	54.4	51.9	63.3	53.2
200 Total		49.6	46.1	44.2	44.9	44.3	44.9	44.2	45.1	49.6	46.1
	10	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7
	20	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9
	50	35.4	50.6	45.6	59.5	45.6	59.5	45.6	59.5	35.4	50.6
	100	62.0	65.8	46.8	64.6	46.8	64.6	46.8	64.6	62.0	65.8
500	200	55.7	59.5	53.2	49.4	53.2	49.4	53.2	49.4	55.7	59.5
	400	54.4	64.6	46.8	63.3	46.8	63.3	46.8	63.3	54.4	64.6
	600	46.8	60.8	40.5	60.8	40.5	60.8	40.5	60.8	46.8	60.8
	800	59.5	57.0	58.2	57.0	58.2	57.0	58.2	57.0	59.5	57.0
	1000	59.5	55.7	49.4	49.4	49.4	48.1	49.4	49.4	59.5	55.7
	1500	54.4	60.8	44.3	58.2	44.3	58.2	44.3	58.2	54.4	60.8
500 Total		45.3	50.1	41.0	48.9	41.0	48.7	41.0	48.9	45.3	50.1
	10	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7
	20	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9	13.9
	50	35.4	44.3	39.2	50.6	39.2	50.6	39.2	50.6	35.4	44.3
1000	100	51.9	57.0	49.4	58.2	49.4	58.2	49.4	58.2	51.9	57.0
	200	57.0	57.0	57.0	53.2	57.0	53.2	57.0	53.2	57.0	57.0
	400	59.5	58.2	59.5	49.4	59.5	49.4	59.5	49.4	59.5	58.2
	600	55.7	53.2	57.0	54.4	57.0	54.4	57.0	54.4	55.7	53.2
	800	55.7	55.7	55.7	57.0	55.7	57.0	55.7	57.0	55.7	55.7

	1000	59.5	59.5	50.6	53.2	50.6	53.2	50.6	53.2	59.5	59.5
	1500	53.2	58.2	48.1	64.6	48.1	64.6	48.1	64.6	53.2	58.2
1000 Total		45.3	47.0	44.2	46.7	44.2	46.7	44.2	46.7	45.3	47.0
2000	10	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7	11.4	12.7
	20	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	50	31.6	34.2	32.9	34.2	32.9	34.2	32.9	34.2	31.6	34.2
	100	59.5	53.2	58.2	57.0	58.2	57.0	58.2	57.0	58.2	53.2
	200	50.6	62.0	41.8	57.0	41.8	57.0	41.8	57.0	50.6	62.0
	400	55.7	62.0	51.9	50.6	51.9	50.6	51.9	50.6	55.7	62.0
	600	41.8	45.6	40.5	50.6	40.5	50.6	40.5	50.6	41.8	45.6
	800	49.4	54.4	51.9	59.5	51.9	58.2	51.9	59.5	49.4	54.4
	1000	50.6	59.5	49.4	55.7	49.4	55.7	49.4	55.7	50.6	59.5
	1500	51.9	57.0	46.8	55.7	46.8	55.7	46.8	55.7	51.9	57.0
2000 Total		41.5	45.3	39.7	44.6	39.7	44.4	39.7	44.6	41.4	45.3

Table 74. Results from the comparison of the Pearson correlation versus the MI calculated using the histogram-based method with fixed number of bins and discrete empirical distribution for detecting the input with the least impact on the output, per number of bins, number of replications, and normalization method.

Bins	**Number of replications**	**MI [%]**		NMI_{arith} **[%]**		NMI_{geom} **[%]**		NMI_{joint} **[%]**		NMI_{theor} **[%]**	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
2	10	3.8	7.6	3.8	7.6	3.8	7.6	3.8	7.6	3.8	7.6
	20	63.3	68.4	63.3	68.4	63.3	68.4	63.3	68.4	63.3	68.4
	50	72.2	72.2	72.2	72.2	72.2	72.2	72.2	72.2	72.2	72.2
	100	86.1	93.7	88.6	92.4	88.6	92.4	88.6	92.4	86.1	93.7
	200	83.5	87.3	82.3	87.3	82.3	87.3	82.3	87.3	83.5	87.3
	400	65.8	78.5	64.6	78.5	64.6	78.5	64.6	78.5	65.8	78.5
	600	68.4	77.2	68.4	77.2	68.4	77.2	68.4	77.2	68.4	77.2
	800	74.7	77.2	74.7	77.2	74.7	77.2	74.7	77.2	74.7	77.2
	1000	75.9	75.9	75.9	75.9	75.9	75.9	75.9	75.9	75.9	75.9
	1500	77.2	78.5	77.2	78.5	77.2	78.5	77.2	78.5	77.2	78.5
2 Total		67.1	71.6	67.1	71.5	67.1	71.5	67.1	71.5	67.1	71.6
5	10	40.5	55.7	45.6	59.5	45.6	59.5	45.6	59.5	39.2	54.4
	20	70.9	73.4	70.9	73.4	70.9	73.4	70.9	73.4	70.9	73.4
	50	72.2	70.9	73.4	70.9	73.4	70.9	73.4	70.9	72.2	70.9
	100	79.7	81.0	78.5	79.7	78.5	79.7	78.5	79.7	79.7	81.0
	200	75.9	82.3	75.9	81.0	75.9	81.0	75.9	81.0	75.9	82.3
	400	64.6	77.2	63.3	75.9	63.3	75.9	63.3	75.9	64.6	77.2

	600	67.1	75.9	67.1	74.7	67.1	74.7	67.1	74.7	67.1	75.9
	800	73.4	78.5	74.7	75.9	74.7	74.7	74.7	75.9	73.4	78.5
	1000	72.2	74.7	70.9	74.7	70.9	74.7	70.9	74.7	72.2	74.7
	1500	67.1	69.6	68.4	70.9	68.4	69.6	68.4	70.9	67.1	69.6
5 Total		68.4	73.9	68.9	73.7	68.9	73.4	68.9	73.7	68.2	73.8
10	10	30.4	46.8	30.4	48.1	30.4	48.1	30.4	48.1	30.4	46.8
	20	55.7	67.1	59.5	69.6	59.5	69.6	59.5	69.6	55.7	67.1
	50	63.3	70.9	64.6	73.4	64.6	73.4	64.6	73.4	63.3	70.9
	100	63.3	74.7	62.0	70.9	62.0	70.9	62.0	70.9	63.3	74.7
	200	54.4	72.2	55.7	68.4	55.7	68.4	55.7	68.4	54.4	72.2
	400	54.4	69.6	57.0	69.6	57.0	69.6	57.0	69.6	54.4	69.6
	600	57.0	67.1	55.7	67.1	55.7	67.1	55.7	67.1	57.0	67.1
	800	58.2	64.6	57.0	67.1	58.2	67.1	57.0	67.1	58.2	64.6
	1000	62.0	67.1	59.5	64.6	59.5	64.6	59.5	64.6	62.0	67.1
	1500	58.2	67.1	55.7	67.1	55.7	67.1	55.7	67.1	58.2	67.1
10 Total		55.7	66.7	55.7	66.6	55.8	66.6	55.7	66.6	55.7	66.7
25	10	3.8	13.9	3.8	13.9	3.8	13.9	3.8	13.9	3.8	13.9
	20	59.5	65.8	57.0	69.6	57.0	69.6	57.0	69.6	59.5	65.8
	50	62.0	74.7	62.0	75.9	62.0	75.9	62.0	75.9	62.0	74.7
	100	54.4	57.0	49.4	58.2	49.4	58.2	49.4	58.2	54.4	57.0
	200	60.8	62.0	59.5	63.3	58.2	63.3	59.5	63.3	60.8	62.0
	400	55.7	59.5	53.2	57.0	53.2	57.0	53.2	57.0	55.7	59.5
	600	57.0	55.7	54.4	55.7	54.4	55.7	54.4	55.7	57.0	55.7
	800	58.2	64.6	57.0	59.5	57.0	59.5	57.0	59.5	58.2	64.6
	1000	55.7	68.4	55.7	64.6	55.7	64.6	55.7	64.6	55.7	68.4
	1500	58.2	62.0	55.7	60.8	55.7	60.8	55.7	60.8	58.2	62.0
25 Total		52.5	58.4	50.8	57.8	50.6	57.8	50.8	57.8	52.5	58.4
50	10	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3
	20	31.6	39.2	38.0	58.2	38.0	58.2	38.0	58.2	31.6	39.2
	50	54.4	62.0	50.6	70.9	50.6	70.9	50.6	70.9	54.4	62.0
	100	65.8	64.6	58.2	59.5	58.2	58.2	58.2	59.5	65.8	64.6
	200	67.1	67.1	60.8	64.6	60.8	64.6	60.8	64.6	67.1	67.1
	400	58.2	60.8	57.0	60.8	57.0	60.8	57.0	60.8	58.2	60.8
	600	59.5	58.2	57.0	59.5	57.0	59.5	57.0	59.5	59.5	58.2
	800	53.2	57.0	55.7	57.0	54.4	57.0	55.7	57.0	53.2	57.0
	1000	55.7	62.0	57.0	60.8	57.0	60.8	57.0	60.8	55.7	62.0
	1500	49.4	60.8	51.9	60.8	51.9	60.8	51.9	60.8	49.4	60.8
50 Total		49.6	53.8	48.7	55.8	48.6	55.7	48.7	55.8	49.6	53.8
100	10	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9

	20	30.4	34.2	35.4	38.0	35.4	38.0	35.4	38.0	30.4	34.2
	50	54.4	54.4	51.9	57.0	51.9	57.0	51.9	57.0	54.4	54.4
	100	67.1	67.1	53.2	62.0	53.2	62.0	53.2	62.0	67.1	67.1
	200	67.1	58.2	64.6	60.8	64.6	60.8	64.6	60.8	67.1	58.2
	400	60.8	57.0	55.7	51.9	55.7	51.9	55.7	51.9	60.8	57.0
	600	62.0	57.0	57.0	54.4	57.0	54.4	57.0	54.4	62.0	57.0
	800	64.6	57.0	63.3	54.4	63.3	54.4	63.3	54.4	64.6	57.0
	1000	67.1	53.2	62.0	54.4	62.0	53.2	62.0	54.4	67.1	53.2
	1500	68.4	54.4	64.6	54.4	64.6	54.4	64.6	54.4	68.4	54.4
100 Total		54.6	50.1	51.1	49.6	51.1	49.5	51.1	49.6	54.6	50.1
	10	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9
	20	10.1	15.2	10.1	15.2	10.1	15.2	10.1	15.2	10.1	15.2
	50	54.4	48.1	55.7	50.6	55.7	51.9	55.7	50.6	54.4	48.1
	100	62.0	58.2	51.9	57.0	53.2	57.0	51.9	57.0	62.0	58.2
200	200	57.0	59.5	55.7	58.2	55.7	58.2	55.7	58.2	57.0	59.5
	400	59.5	60.8	57.0	57.0	57.0	57.0	57.0	57.0	59.5	60.8
	600	59.5	60.8	58.2	63.3	58.2	63.3	58.2	63.3	59.5	60.8
	800	68.4	58.2	57.0	55.7	57.0	55.7	57.0	55.7	68.4	58.2
	1000	63.3	59.5	57.0	54.4	57.0	54.4	57.0	54.4	63.3	59.5
	1500	72.2	57.0	63.3	55.7	63.3	55.7	63.3	55.7	72.2	57.0
200 Total		51.0	48.6	47.0	47.6	47.1	47.7	47.0	47.6	51.0	48.6
	10	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3
	20	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9
	50	36.7	54.4	48.1	60.8	48.1	60.8	48.1	60.8	36.7	54.4
	100	60.8	65.8	48.1	63.3	48.1	63.3	48.1	63.3	60.8	65.8
500	200	58.2	63.3	58.2	55.7	58.2	55.7	58.2	55.7	58.2	63.3
	400	59.5	63.3	50.6	62.0	50.6	62.0	50.6	62.0	59.5	63.3
	600	54.4	63.3	48.1	62.0	48.1	62.0	48.1	62.0	54.4	63.3
	800	69.6	62.0	64.6	67.1	64.6	67.1	64.6	67.1	69.6	62.0
	1000	68.4	63.3	55.7	58.2	55.7	57.0	55.7	58.2	68.4	63.3
	1500	60.8	64.6	49.4	65.8	49.4	65.8	49.4	65.8	60.8	64.6
500 Total		47.3	51.5	42.8	51.0	42.8	50.9	42.8	51.0	47.3	51.5
	10	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3
	20	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9
	50	38.0	54.4	41.8	55.7	41.8	55.7	41.8	55.7	38.0	54.4
1000	100	53.2	55.7	50.6	54.4	50.6	54.4	50.6	54.4	53.2	55.7
	200	54.4	58.2	55.7	54.4	55.7	54.4	55.7	54.4	54.4	58.2
	400	62.0	63.3	65.8	51.9	65.8	51.9	65.8	51.9	62.0	63.3
	600	59.5	58.2	60.8	60.8	60.8	60.8	60.8	60.8	59.5	58.2

	800	64.6	65.8	69.6	62.0	69.6	62.0	69.6	62.0	64.6	65.8
	1000	63.3	60.8	58.2	59.5	58.2	59.5	58.2	59.5	63.3	60.8
	1500	58.2	64.6	53.2	65.8	53.2	65.8	53.2	65.8	58.2	64.6
1000 Total		45.8	49.6	46.1	48.0	46.1	48.0	46.1	48.0	45.8	49.6
2000	10	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3	1.3	6.3
	20	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9	3.8	8.9
	50	30.4	34.2	31.6	32.9	31.6	32.9	31.6	32.9	30.4	34.2
	100	60.8	53.2	59.5	57.0	59.5	57.0	59.5	57.0	59.5	53.2
	200	53.2	64.6	46.8	62.0	46.8	62.0	46.8	62.0	53.2	65.8
	400	59.5	65.8	57.0	54.4	57.0	54.4	57.0	54.4	59.5	65.8
	600	51.9	55.7	49.4	57.0	49.4	57.0	49.4	57.0	51.9	55.7
	800	60.8	60.8	59.5	59.5	59.5	58.2	59.5	59.5	60.8	60.8
	1000	58.2	60.8	57.0	58.2	57.0	58.2	57.0	58.2	58.2	60.8
	1500	58.2	60.8	50.6	62.0	50.6	62.0	50.6	62.0	58.2	60.8
2000 Total		43.8	47.1	41.6	45.8	41.6	45.7	41.6	45.8	43.7	47.2

Table 75. Results from the comparison of the R^2_{adj} versus the MI calculated using the histogram-based method with fixed number of bins and discrete empirical distribution for detecting the input with the greatest impact on the output, per number of bins, number of replications, and normalization method.

Bins	**Number of replications**	**MI [%]**		NMI_{arith} **[%]**		NMI_{geom} **[%]**		NMI_{joint} **[%]**		NMI_{theor} **[%]**	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
2	10	1.3	11.4	1.3	11.4	1.3	11.4	1.3	11.4	1.3	11.4
	20	60.8	67.1	60.8	67.1	60.8	67.1	60.8	67.1	60.8	67.1
	50	68.4	70.9	68.4	70.9	68.4	70.9	68.4	70.9	68.4	70.9
	100	75.9	82.3	78.5	81.0	78.5	81.0	78.5	81.0	75.9	82.3
	200	72.2	79.7	69.6	79.7	69.6	79.7	69.6	79.7	72.2	79.7
	400	58.2	75.9	58.2	75.9	58.2	75.9	58.2	75.9	58.2	75.9
	600	58.2	74.7	59.5	74.7	59.5	74.7	59.5	74.7	58.2	74.7
	800	63.3	70.9	63.3	70.9	63.3	70.9	63.3	70.9	63.3	70.9
	1000	67.1	70.9	67.1	70.9	67.1	70.9	67.1	70.9	67.1	70.9
	1500	68.4	74.7	68.4	74.7	68.4	74.7	68.4	74.7	68.4	74.7
2 Total		59.4	67.8	59.5	67.7	59.5	67.7	59.5	67.7	59.4	67.8
5	10	41.8	53.2	48.1	57.0	48.1	57.0	48.1	57.0	40.5	53.2
	20	70.9	69.6	69.6	69.6	69.6	69.6	69.6	69.6	70.9	69.6
	50	69.6	67.1	69.6	64.6	69.6	64.6	69.6	64.6	69.6	67.1
	100	72.2	77.2	70.9	75.9	70.9	75.9	70.9	75.9	72.2	77.2

	200	69.6	75.9	67.1	74.7	67.1	74.7	67.1	74.7	69.6	75.9
	400	57.0	73.4	55.7	72.2	55.7	72.2	55.7	72.2	57.0	73.4
	600	55.7	72.2	55.7	70.9	55.7	70.9	55.7	70.9	55.7	72.2
	800	63.3	70.9	64.6	70.9	64.6	69.6	64.6	70.9	63.3	70.9
	1000	64.6	69.6	65.8	72.2	65.8	72.2	65.8	72.2	64.6	69.6
	1500	58.2	67.1	60.8	67.1	60.8	65.8	60.8	67.1	58.2	67.1
5 Total		62.3	69.6	62.8	69.5	62.8	69.2	62.8	69.5	62.2	69.6
10	10	30.4	48.1	30.4	49.4	30.4	49.4	30.4	49.4	30.4	48.1
	20	55.7	74.7	58.2	74.7	58.2	74.7	58.2	74.7	55.7	74.7
	50	57.0	67.1	55.7	68.4	55.7	68.4	55.7	68.4	57.0	67.1
	100	55.7	65.8	55.7	62.0	55.7	62.0	55.7	62.0	55.7	65.8
	200	46.8	64.6	44.3	59.5	44.3	58.2	44.3	59.5	46.8	64.6
	400	44.3	63.3	44.3	62.0	44.3	62.0	44.3	62.0	44.3	63.3
	600	44.3	63.3	40.5	60.8	40.5	60.8	40.5	60.8	44.3	63.3
	800	45.6	58.2	43.0	59.5	44.3	59.5	43.0	59.5	45.6	58.2
	1000	46.8	62.0	44.3	58.2	44.3	58.2	44.3	58.2	46.8	62.0
	1500	44.3	64.6	41.8	63.3	41.8	63.3	41.8	63.3	44.3	64.6
10 Total		47.1	63.2	45.8	61.8	45.9	61.6	45.8	61.8	47.1	63.2
25	10	10.1	15.2	11.4	15.2	11.4	15.2	10.1	15.2	10.1	15.2
	20	59.5	67.1	58.2	67.1	58.2	67.1	58.2	67.1	59.5	67.1
	50	55.7	74.7	55.7	77.2	55.7	77.2	55.7	77.2	55.7	74.7
	100	46.8	57.0	44.3	54.4	44.3	54.4	44.3	54.4	46.8	57.0
	200	55.7	62.0	53.2	60.8	51.9	60.8	53.2	60.8	55.7	62.0
	400	49.4	58.2	46.8	54.4	46.8	54.4	46.8	54.4	49.4	58.2
	600	48.1	50.6	45.6	50.6	45.6	50.6	45.6	50.6	48.1	50.6
	800	45.6	57.0	44.3	54.4	44.3	54.4	44.3	54.4	45.6	57.0
	1000	44.3	65.8	45.6	62.0	45.6	62.0	45.6	62.0	44.3	65.8
	1500	51.9	60.8	50.6	59.5	50.6	59.5	50.6	59.5	51.9	60.8
25 Total		46.7	56.8	45.6	55.6	45.4	55.6	45.4	55.6	46.7	56.8
50	10	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1
	20	32.9	38.0	38.0	55.7	38.0	55.7	38.0	55.7	32.9	38.0
	50	48.1	54.4	48.1	63.3	48.1	63.3	48.1	63.3	48.1	54.4
	100	62.0	62.0	53.2	59.5	53.2	59.5	53.2	59.5	62.0	62.0
	200	64.6	65.8	55.7	65.8	55.7	65.8	55.7	65.8	64.6	65.8
	400	58.2	58.2	57.0	59.5	57.0	59.5	57.0	59.5	58.2	58.2
	600	59.5	59.5	57.0	57.0	57.0	57.0	57.0	57.0	59.5	59.5
	800	40.5	53.2	44.3	51.9	43.0	51.9	44.3	51.9	40.5	53.2
	1000	45.6	60.8	45.6	59.5	45.6	59.5	45.6	59.5	45.6	60.8
	1500	44.3	55.7	46.8	55.7	46.8	55.7	46.8	55.7	44.3	55.7

50 Total		46.1	51.3	45.1	53.3	44.9	53.3	45.1	53.3	46.1	51.3
100	10	7.6	8.9	7.6	8.9	7.6	8.9	7.6	8.9	7.6	8.9
	20	31.6	36.7	36.7	40.5	36.7	40.5	36.7	40.5	31.6	36.7
	50	49.4	51.9	46.8	55.7	46.8	55.7	46.8	55.7	49.4	51.9
	100	64.6	65.8	51.9	59.5	51.9	60.8	51.9	59.5	64.6	65.8
	200	59.5	57.0	55.7	59.5	55.7	59.5	55.7	59.5	59.5	57.0
	400	53.2	50.6	49.4	46.8	50.6	46.8	49.4	46.8	53.2	50.6
	600	55.7	49.4	50.6	46.8	50.6	45.6	50.6	46.8	55.7	49.4
	800	53.2	53.2	51.9	50.6	51.9	50.6	51.9	50.6	53.2	53.2
	1000	62.0	51.9	58.2	50.6	58.2	49.4	58.2	50.6	62.0	51.9
	1500	59.5	51.9	57.0	45.6	57.0	45.6	57.0	45.6	59.5	51.9
100 Total		49.6	47.7	46.6	46.5	46.7	46.3	46.6	46.5	49.6	47.7
200	10	10.1	10.1	10.1	10.1	10.1	10.1	10.1	10.1	10.1	10.1
	20	10.1	16.5	10.1	16.5	10.1	16.5	10.1	16.5	10.1	16.5
	50	57.0	45.6	55.7	48.1	55.7	48.1	55.7	49.4	57.0	45.6
	100	57.0	54.4	46.8	55.7	48.1	55.7	46.8	55.7	57.0	54.4
	200	55.7	58.2	53.2	51.9	53.2	51.9	53.2	51.9	55.7	58.2
	400	55.7	55.7	53.2	51.9	53.2	51.9	53.2	51.9	55.7	55.7
	600	53.2	54.4	49.4	57.0	49.4	57.0	49.4	57.0	53.2	54.4
	800	60.8	51.9	48.1	51.9	48.1	51.9	48.1	51.9	60.8	51.9
	1000	64.6	57.0	51.9	50.6	51.9	50.6	51.9	50.6	64.6	57.0
	1500	63.3	53.2	54.4	51.9	54.4	51.9	54.4	51.9	63.3	53.2
200 Total		48.7	45.7	43.3	44.6	43.4	44.6	43.3	44.7	48.7	45.7
500	10	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9
	20	7.6	13.9	7.6	13.9	7.6	13.9	7.6	13.9	7.6	13.9
	50	35.4	50.6	45.6	59.5	45.6	59.5	45.6	59.5	35.4	50.6
	100	62.0	65.8	46.8	64.6	46.8	64.6	46.8	64.6	62.0	65.8
	200	55.7	59.5	53.2	49.4	53.2	49.4	53.2	49.4	55.7	59.5
	400	54.4	64.6	46.8	63.3	46.8	63.3	46.8	63.3	54.4	64.6
	600	46.8	60.8	40.5	60.8	40.5	60.8	40.5	60.8	46.8	60.8
	800	59.5	57.0	58.2	57.0	58.2	57.0	58.2	57.0	59.5	57.0
	1000	59.5	55.7	49.4	49.4	49.4	48.1	49.4	49.4	59.5	55.7
	1500	54.4	60.8	44.3	58.2	44.3	58.2	44.3	58.2	54.4	60.8
500 Total		44.4	49.7	40.1	48.5	40.1	48.4	40.1	48.5	44.4	49.7
1000	10	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9
	20	7.6	13.9	7.6	13.9	7.6	13.9	7.6	13.9	7.6	13.9
	50	35.4	44.3	39.2	50.6	39.2	50.6	39.2	50.6	35.4	44.3
	100	51.9	57.0	49.4	58.2	49.4	58.2	49.4	58.2	51.9	57.0
	200	57.0	57.0	57.0	53.2	57.0	53.2	57.0	53.2	57.0	57.0

	400	59.5	58.2	59.5	49.4	59.5	49.4	59.5	49.4	59.5	58.2
	600	55.7	53.2	57.0	54.4	57.0	54.4	57.0	54.4	55.7	53.2
	800	55.7	55.7	55.7	57.0	55.7	57.0	55.7	57.0	55.7	55.7
	1000	59.5	59.5	50.6	53.2	50.6	53.2	50.6	53.2	59.5	59.5
	1500	53.2	58.2	48.1	64.6	48.1	64.6	48.1	64.6	53.2	58.2
1000 Total		44.4	46.6	43.3	46.3	43.3	46.3	43.3	46.3	44.4	46.6
2000	10	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9	8.9
	20	6.3	12.7	6.3	12.7	6.3	12.7	6.3	12.7	6.3	12.7
	50	31.6	34.2	32.9	34.2	32.9	34.2	32.9	34.2	31.6	34.2
	100	59.5	53.2	58.2	57.0	58.2	57.0	58.2	57.0	58.2	53.2
	200	50.6	62.0	41.8	57.0	41.8	57.0	41.8	57.0	50.6	62.0
	400	55.7	62.0	51.9	50.6	51.9	50.6	51.9	50.6	55.7	62.0
	600	41.8	45.6	40.5	50.6	40.5	50.6	40.5	50.6	41.8	45.6
	800	49.4	54.4	51.9	59.5	51.9	58.2	51.9	59.5	49.4	54.4
	1000	50.6	59.5	49.4	55.7	49.4	55.7	49.4	55.7	50.6	59.5
	1500	51.9	57.0	46.8	55.7	46.8	55.7	46.8	55.7	51.9	57.0
2000 Total		40.6	44.9	38.9	44.2	38.9	44.1	38.9	44.2	40.5	44.9

Table 76. Results from the comparison of the R^2_{adj} versus the MI calculated using the histogram-based method with fixed number of bins and discrete empirical distribution for detecting the input with the least impact on the output, per number of bins, number of replications, and normalization method.

Bins	**Number of replications**	**MI [%]**		NMI_{arith} **[%]**		NMI_{geom} **[%]**		NMI_{joint} **[%]**		NMI_{theor} **[%]**	
		NIS	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**	**NIS**	**TIS**
2	10	2.5	1.3	2.5	1.3	2.5	1.3	2.5	1.3	2.5	1.3
	20	62.0	63.3	62.0	63.3	62.0	63.3	62.0	63.3	62.0	63.3
	50	68.4	72.2	68.4	72.2	68.4	72.2	68.4	72.2	68.4	72.2
	100	83.5	93.7	86.1	92.4	86.1	92.4	86.1	92.4	83.5	93.7
	200	83.5	87.3	82.3	87.3	82.3	87.3	82.3	87.3	83.5	87.3
	400	65.8	78.5	64.6	78.5	64.6	78.5	64.6	78.5	65.8	78.5
	600	68.4	77.2	68.4	77.2	68.4	77.2	68.4	77.2	68.4	77.2
	800	74.7	77.2	74.7	77.2	74.7	77.2	74.7	77.2	74.7	77.2
	1000	75.9	75.9	75.9	75.9	75.9	75.9	75.9	75.9	75.9	75.9
	1500	77.2	78.5	77.2	78.5	77.2	78.5	77.2	78.5	77.2	78.5
2 Total		66.2	70.5	66.2	70.4	66.2	70.4	66.2	70.4	66.2	70.5
5	10	40.5	49.4	45.6	53.2	45.6	53.2	45.6	53.2	39.2	48.1
	20	69.6	68.4	69.6	68.4	69.6	68.4	69.6	68.4	69.6	68.4

	50	68.4	70.9	69.6	70.9	69.6	70.9	69.6	70.9	68.4	70.9
	100	77.2	81.0	75.9	79.7	75.9	79.7	75.9	79.7	77.2	81.0
	200	75.9	82.3	75.9	81.0	75.9	81.0	75.9	81.0	75.9	82.3
	400	64.6	77.2	63.3	75.9	63.3	75.9	63.3	75.9	64.6	77.2
	600	67.1	75.9	67.1	74.7	67.1	74.7	67.1	74.7	67.1	75.9
	800	73.4	78.5	74.7	75.9	74.7	74.7	74.7	75.9	73.4	78.5
	1000	72.2	74.7	70.9	74.7	70.9	74.7	70.9	74.7	72.2	74.7
	1500	67.1	69.6	68.4	70.9	68.4	69.6	68.4	70.9	67.1	69.6
5 Total		67.6	72.8	68.1	72.5	68.1	72.3	68.1	72.5	67.5	72.7
10	10	29.1	40.5	29.1	41.8	29.1	41.8	29.1	41.8	29.1	40.5
	20	54.4	62.0	58.2	64.6	58.2	64.6	58.2	64.6	54.4	62.0
	50	59.5	70.9	60.8	73.4	60.8	73.4	60.8	73.4	59.5	70.9
	100	60.8	74.7	59.5	70.9	59.5	70.9	59.5	70.9	60.8	74.7
	200	54.4	72.2	55.7	68.4	55.7	68.4	55.7	68.4	54.4	72.2
	400	54.4	69.6	57.0	69.6	57.0	69.6	57.0	69.6	54.4	69.6
	600	57.0	67.1	55.7	67.1	55.7	67.1	55.7	67.1	57.0	67.1
	800	58.2	64.6	57.0	67.1	58.2	67.1	57.0	67.1	58.2	64.6
	1000	62.0	67.1	59.5	64.6	59.5	64.6	59.5	64.6	62.0	67.1
	1500	58.2	67.1	55.7	67.1	55.7	67.1	55.7	67.1	58.2	67.1
10 Total		54.8	65.6	54.8	65.4	54.9	65.4	54.8	65.4	54.8	65.6
25	10	3.8	7.6	3.8	7.6	3.8	7.6	3.8	7.6	3.8	7.6
	20	58.2	60.8	55.7	64.6	55.7	64.6	55.7	64.6	58.2	60.8
	50	58.2	74.7	58.2	75.9	58.2	75.9	58.2	75.9	58.2	74.7
	100	51.9	57.0	46.8	58.2	46.8	58.2	46.8	58.2	51.9	57.0
	200	60.8	62.0	59.5	63.3	58.2	63.3	59.5	63.3	60.8	62.0
	400	55.7	59.5	53.2	57.0	53.2	57.0	53.2	57.0	55.7	59.5
	600	57.0	55.7	54.4	55.7	54.4	55.7	54.4	55.7	57.0	55.7
	800	58.2	64.6	57.0	59.5	57.0	59.5	57.0	59.5	58.2	64.6
	1000	55.7	68.4	55.7	64.6	55.7	64.6	55.7	64.6	55.7	68.4
	1500	58.2	62.0	55.7	60.8	55.7	60.8	55.7	60.8	58.2	62.0
25 Total		51.8	57.2	50.0	56.7	49.9	56.7	50.0	56.7	51.8	57.2
50	10	2.5	0.0	2.5	0.0	2.5	0.0	2.5	0.0	2.5	0.0
	20	30.4	34.2	36.7	53.2	36.7	53.2	36.7	53.2	30.4	34.2
	50	50.6	62.0	46.8	70.9	46.8	70.9	46.8	70.9	50.6	62.0
	100	63.3	64.6	55.7	59.5	55.7	58.2	55.7	59.5	63.3	64.6
	200	67.1	67.1	60.8	64.6	60.8	64.6	60.8	64.6	67.1	67.1
	400	58.2	60.8	57.0	60.8	57.0	60.8	57.0	60.8	58.2	60.8
	600	59.5	58.2	57.0	59.5	57.0	59.5	57.0	59.5	59.5	58.2
	800	53.2	57.0	55.7	57.0	54.4	57.0	55.7	57.0	53.2	57.0

	1000	55.7	62.0	57.0	60.8	57.0	60.8	57.0	60.8	55.7	62.0
	1500	49.4	60.8	51.9	60.8	51.9	60.8	51.9	60.8	49.4	60.8
50 Total		49.0	52.7	48.1	54.7	48.0	54.6	48.1	54.7	49.0	52.7
100	10	3.8	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8	2.5
	20	29.1	29.1	34.2	32.9	34.2	32.9	34.2	32.9	29.1	29.1
	50	50.6	54.4	48.1	57.0	48.1	57.0	48.1	57.0	50.6	54.4
	100	64.6	67.1	50.6	62.0	50.6	62.0	50.6	62.0	64.6	67.1
	200	67.1	58.2	64.6	60.8	64.6	60.8	64.6	60.8	67.1	58.2
	400	60.8	57.0	55.7	51.9	55.7	51.9	55.7	51.9	60.8	57.0
	600	62.0	57.0	57.0	54.4	57.0	54.4	57.0	54.4	62.0	57.0
	800	64.6	57.0	63.3	54.4	63.3	54.4	63.3	54.4	64.6	57.0
	1000	67.1	53.2	62.0	54.4	62.0	53.2	62.0	54.4	67.1	53.2
	1500	68.4	54.4	64.6	54.4	64.6	54.4	64.6	54.4	68.4	54.4
100 Total		53.8	49.0	50.4	48.5	50.4	48.4	50.4	48.5	53.8	49.0
200	10	3.8	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8	2.5
	20	8.9	10.1	8.9	10.1	8.9	10.1	8.9	10.1	8.9	10.1
	50	50.6	48.1	51.9	50.6	51.9	51.9	51.9	50.6	50.6	48.1
	100	59.5	58.2	49.4	57.0	50.6	57.0	49.4	57.0	59.5	58.2
	200	57.0	59.5	55.7	58.2	55.7	58.2	55.7	58.2	57.0	59.5
	400	59.5	60.8	57.0	57.0	57.0	57.0	57.0	57.0	59.5	60.8
	600	59.5	60.8	58.2	63.3	58.2	63.3	58.2	63.3	59.5	60.8
	800	68.4	58.2	57.0	55.7	57.0	55.7	57.0	55.7	68.4	58.2
	1000	63.3	59.5	57.0	54.4	57.0	54.4	57.0	54.4	63.3	59.5
	1500	72.2	57.0	63.3	55.7	63.3	55.7	63.3	55.7	72.2	57.0
200 Total		50.3	47.5	46.2	46.5	46.3	46.6	46.2	46.5	50.3	47.5
500	10	1.3	0.0	1.3	0.0	1.3	0.0	1.3	0.0	1.3	0.0
	20	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8
	50	32.9	54.4	44.3	60.8	44.3	60.8	44.3	60.8	32.9	54.4
	100	58.2	65.8	45.6	63.3	45.6	63.3	45.6	63.3	58.2	65.8
	200	58.2	63.3	58.2	55.7	58.2	55.7	58.2	55.7	58.2	63.3
	400	59.5	63.3	50.6	62.0	50.6	62.0	50.6	62.0	59.5	63.3
	600	54.4	63.3	48.1	62.0	48.1	62.0	48.1	62.0	54.4	63.3
	800	69.6	62.0	64.6	67.1	64.6	67.1	64.6	67.1	69.6	62.0
	1000	68.4	63.3	55.7	58.2	55.7	57.0	55.7	58.2	68.4	63.3
	1500	60.8	64.6	49.4	65.8	49.4	65.8	49.4	65.8	60.8	64.6
500 Total		46.6	50.4	42.0	49.9	42.0	49.7	42.0	49.9	46.6	50.4
1000	10	1.3	0.0	1.3	0.0	1.3	0.0	1.3	0.0	1.3	0.0
	20	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8
	50	34.2	54.4	38.0	55.7	38.0	55.7	38.0	55.7	34.2	54.4

	100	50.6	55.7	48.1	54.4	48.1	54.4	48.1	54.4	50.6	55.7
	200	54.4	58.2	55.7	54.4	55.7	54.4	55.7	54.4	54.4	58.2
	400	62.0	63.3	65.8	51.9	65.8	51.9	65.8	51.9	62.0	63.3
	600	59.5	58.2	60.8	60.8	60.8	60.8	60.8	60.8	59.5	58.2
	800	64.6	65.8	69.6	62.0	69.6	62.0	69.6	62.0	64.6	65.8
	1000	63.3	60.8	58.2	59.5	58.2	59.5	58.2	59.5	63.3	60.8
	1500	58.2	64.6	53.2	65.8	53.2	65.8	53.2	65.8	58.2	64.6
1000 Total		45.1	48.5	45.3	46.8	45.3	46.8	45.3	46.8	45.1	48.5
	10	1.3	0.0	1.3	0.0	1.3	0.0	1.3	0.0	1.3	0.0
	20	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8	2.5	3.8
	50	26.6	34.2	27.8	32.9	27.8	32.9	27.8	32.9	26.6	34.2
	100	58.2	53.2	57.0	57.0	57.0	57.0	57.0	57.0	57.0	53.2
	200	53.2	64.6	46.8	62.0	46.8	62.0	46.8	62.0	53.2	65.8
2000	400	59.5	65.8	57.0	54.4	57.0	54.4	57.0	54.4	59.5	65.8
	600	51.9	55.7	49.4	57.0	49.4	57.0	49.4	57.0	51.9	55.7
	800	60.8	60.8	59.5	59.5	59.5	58.2	59.5	59.5	60.8	60.8
	1000	58.2	60.8	57.0	58.2	57.0	58.2	57.0	58.2	58.2	60.8
	1500	58.2	60.8	50.6	62.0	50.6	62.0	50.6	62.0	58.2	60.8
2000 Total		43.0	45.9	40.9	44.7	40.9	44.6	40.9	44.7	42.9	46.1

APPENDIX C

RESULTS OF SECTION 3

Table 77. Results from the comparison of the measures of dependence versus the MI calculated using the kernel method for detecting the input with the greatest impact on the output per kernel function, per bandwidth value, and per number of replications.

Bandwidth type	Kernel function	Bandwidth value	Number of replications	Distance correlation		Pearson correlation		R^2_{adj}	
				NIS %	TIS %	NIS %	TIS %	NIS %	TIS %
Silverman	Epanechnikov	NA	10	52.21	52.21	48.67	49.56	50.44	52.21
			20	41.59	46.90	40.71	46.90	44.25	46.90
			50	38.94	47.79	38.05	47.79	38.05	47.79
			100	40.71	42.48	32.74	38.94	32.74	38.94
			200	43.36	41.59	29.20	36.28	29.20	36.28
			400	51.33	52.21	39.82	48.67	39.82	48.67
			600	53.98	52.21	40.71	53.98	40.71	53.98
			800	44.25	46.90	32.74	46.90	32.74	46.90
			1000	41.59	46.02	34.51	46.02	34.51	46.02
			1500	41.59	46.90	36.28	46.90	36.28	46.90
	Normal	NA	10	47.79	48.67	46.02	46.02	47.79	48.67
			20	38.05	44.25	37.17	45.13	40.71	45.13
			50	39.82	46.02	37.17	46.02	37.17	46.02
			100	39.82	40.71	31.86	37.17	31.86	37.17
			200	48.67	46.02	34.51	40.71	34.51	40.71
			400	53.98	53.98	42.48	50.44	42.48	50.44
			600	53.98	53.10	40.71	54.87	40.71	54.87
			800	44.25	47.79	32.74	47.79	32.74	47.79
			1000	41.59	46.90	34.51	46.90	34.51	46.90
			1500	41.59	46.90	36.28	46.90	36.28	46.90
Trial	Epanechnikov	0.1	10	39.82	41.59	40.71	38.94	42.48	41.59
			20	38.94	44.25	43.36	46.02	46.90	46.02
			50	37.17	46.02	36.28	46.02	36.28	46.02
			100	39.82	39.82	31.86	36.28	31.86	36.28
			200	44.25	43.36	31.86	38.05	31.86	38.05
			400	52.21	53.10	40.71	49.56	40.71	49.56
			600	53.98	53.10	40.71	54.87	40.71	54.87
			800	45.13	46.90	31.86	46.90	31.86	46.90

	1000	42.48	46.02	33.63	46.02	33.63	46.02
	1500	41.59	46.90	36.28	46.90	36.28	46.90
0.2	10	38.94	43.36	42.48	40.71	44.25	43.36
	20	39.82	44.25	38.94	43.36	42.48	43.36
	50	38.05	46.90	37.17	46.90	37.17	46.90
	100	41.59	41.59	31.86	38.05	31.86	38.05
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	51.33	52.21	39.82	48.67	39.82	48.67
	600	53.10	52.21	39.82	53.98	39.82	53.98
	800	45.13	46.90	31.86	46.90	31.86	46.90
	1000	42.48	46.02	33.63	46.02	33.63	46.02
	1500	41.59	47.79	36.28	47.79	36.28	47.79
0.5	10	50.44	52.21	46.90	49.56	48.67	52.21
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.94	47.79	38.05	47.79	38.05	47.79
	100	41.59	42.48	31.86	38.94	31.86	38.94
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	53.10	53.10	41.59	49.56	41.59	49.56
	600	54.87	53.98	41.59	55.75	41.59	55.75
	800	41.59	44.25	28.32	42.48	28.32	42.48
	1000	41.59	44.25	30.09	42.48	30.09	42.48
	1500	42.48	47.79	37.17	47.79	37.17	47.79
0.01	10	7.96	0.88	8.85	0.88	10.62	3.54
	20	43.36	38.05	40.71	39.82	44.25	39.82
	50	37.17	45.13	40.71	50.44	40.71	50.44
	100	38.05	38.94	30.09	33.63	30.09	33.63
	200	46.02	41.59	30.09	34.51	30.09	34.51
	400	54.87	53.10	41.59	49.56	41.59	49.56
	600	55.75	54.87	42.48	56.64	42.48	56.64
	800	46.02	37.17	32.74	37.17	32.74	37.17
	1000	44.25	37.17	32.74	37.17	32.74	37.17
	1500	42.48	36.28	37.17	36.28	37.17	36.28
0.001	10	5.31	0.00	6.19	0.00	7.96	2.65
	20	4.42	0.88	4.42	0.88	7.96	0.88
	50	39.82	43.36	38.94	41.59	38.94	41.59
	100	50.44	46.02	38.94	47.79	38.94	47.79
	200	47.79	34.51	43.36	38.94	43.36	38.94
	400	50.44	39.82	37.17	36.28	37.17	36.28
	600	52.21	36.28	38.94	38.05	38.94	38.05

	800	44.25	30.09	30.97	30.09	30.97	30.09
	1000	39.82	28.32	29.20	28.32	29.20	28.32
	1500	38.94	27.43	33.63	29.20	33.63	29.20
0.0001	10	5.31	0.00	6.19	0.00	7.96	2.65
	20	5.31	0.00	5.31	0.00	8.85	0.00
	50	11.50	11.50	9.73	11.50	9.73	11.50
	100	39.82	37.17	45.13	46.90	45.13	46.90
	200	43.36	33.63	45.13	38.94	45.13	38.94
	400	42.48	38.94	43.36	37.17	43.36	37.17
	600	41.59	40.71	42.48	42.48	42.48	42.48
	800	41.59	32.74	35.40	32.74	35.40	32.74
	1000	37.17	30.97	34.51	30.09	34.51	30.09
	1500	43.36	34.51	39.82	33.63	39.82	33.63
1	10	50.44	52.21	46.90	49.56	48.67	52.21
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.94	47.79	38.05	47.79	38.05	47.79
	100	41.59	41.59	31.86	38.05	31.86	38.05
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	53.10	53.10	41.59	49.56	41.59	49.56
	600	54.87	53.98	41.59	55.75	41.59	55.75
	800	42.48	45.13	29.20	45.13	29.20	45.13
	1000	43.36	46.90	33.63	46.90	33.63	46.90
	1500	42.48	48.67	37.17	48.67	37.17	48.67
1.5	10	50.44	52.21	46.90	49.56	48.67	52.21
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.94	47.79	38.05	47.79	38.05	47.79
	100	41.59	41.59	31.86	38.05	31.86	38.05
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	53.10	53.10	41.59	49.56	41.59	49.56
	600	54.87	53.98	41.59	55.75	41.59	55.75
	800	42.48	45.13	29.20	45.13	29.20	45.13
	1000	43.36	46.90	33.63	46.90	33.63	46.90
	1500	42.48	48.67	37.17	48.67	37.17	48.67
5	10	50.44	52.21	46.90	49.56	48.67	52.21
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.94	47.79	38.05	47.79	38.05	47.79
	100	41.59	41.59	31.86	38.05	31.86	38.05
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	53.10	53.10	41.59	49.56	41.59	49.56

		600	54.87	53.98	41.59	55.75	41.59	55.75
		800	43.36	46.02	28.32	46.02	28.32	46.02
		1000	43.36	46.90	33.63	46.90	33.63	46.90
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	10	10	50.44	52.21	46.90	49.56	48.67	52.21
		20	40.71	46.02	39.82	46.02	43.36	46.02
		50	38.94	47.79	38.05	47.79	38.05	47.79
		100	41.59	41.59	31.86	38.05	31.86	38.05
		200	44.25	43.36	31.86	38.05	31.86	38.05
		400	53.10	53.10	41.59	49.56	41.59	49.56
		600	54.87	53.98	41.59	55.75	41.59	55.75
		800	43.36	46.02	28.32	46.02	28.32	46.02
		1000	43.36	46.90	33.63	46.90	33.63	46.90
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	100	10	50.44	52.21	46.90	49.56	48.67	52.21
		20	40.71	46.02	39.82	46.02	43.36	46.02
		50	38.94	47.79	38.05	47.79	38.05	47.79
		100	41.59	41.59	31.86	38.05	31.86	38.05
		200	44.25	43.36	31.86	38.05	31.86	38.05
		400	53.10	53.10	41.59	49.56	41.59	49.56
		600	54.87	53.98	41.59	55.75	41.59	55.75
		800	43.36	46.02	28.32	46.02	28.32	46.02
		1000	43.36	46.90	33.63	46.90	33.63	46.90
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	1000	10	50.44	52.21	46.90	49.56	48.67	52.21
		20	40.71	46.02	39.82	46.02	43.36	46.02
		50	38.94	47.79	38.05	47.79	38.05	47.79
		100	46.02	44.25	36.28	40.71	36.28	40.71
		200	46.90	40.71	33.63	35.40	33.63	35.40
		400	56.64	53.10	45.13	49.56	45.13	49.56
		600	56.64	53.10	44.25	54.87	44.25	54.87
		800	41.59	44.25	28.32	44.25	28.32	44.25
		1000	42.48	46.90	32.74	46.90	32.74	46.90
		1500	42.48	46.90	37.17	46.90	37.17	46.90
Normal	0.1	10	34.51	35.40	35.40	34.51	37.17	37.17
		20	39.82	40.71	40.71	41.59	44.25	41.59
		50	39.82	46.90	33.63	43.36	33.63	43.36
		100	40.71	40.71	30.97	37.17	30.97	37.17
		200	46.90	46.02	34.51	40.71	34.51	40.71

	400	52.21	53.10	40.71	49.56	40.71	49.56
	600	53.98	53.10	40.71	54.87	40.71	54.87
	800	45.13	46.90	31.86	46.90	31.86	46.90
	1000	42.48	46.90	33.63	46.90	33.63	46.90
	1500	41.59	46.90	36.28	46.90	36.28	46.90
0.2	10	38.05	44.25	41.59	39.82	43.36	42.48
	20	39.82	45.13	38.94	44.25	42.48	44.25
	50	37.17	46.02	36.28	46.02	36.28	46.02
	100	40.71	41.59	30.97	38.05	30.97	38.05
	200	42.48	41.59	30.09	36.28	30.09	36.28
	400	51.33	52.21	39.82	48.67	39.82	48.67
	600	53.98	53.10	40.71	54.87	40.71	54.87
	800	45.13	46.90	31.86	46.90	31.86	46.90
	1000	42.48	46.02	33.63	46.02	33.63	46.02
	1500	41.59	46.90	36.28	46.90	36.28	46.90
0.5	10	49.56	51.33	46.02	48.67	47.79	51.33
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.05	46.90	37.17	46.90	37.17	46.90
	100	41.59	41.59	31.86	38.05	31.86	38.05
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	52.21	52.21	40.71	48.67	40.71	48.67
	600	53.98	53.10	40.71	54.87	40.71	54.87
	800	44.25	46.02	29.20	46.02	29.20	46.02
	1000	41.59	45.13	32.74	45.13	32.74	45.13
	1500	42.48	48.67	37.17	48.67	37.17	48.67
0.01	10	50.44	46.02	47.79	43.36	49.56	46.02
	20	41.59	36.28	37.17	37.17	40.71	37.17
	50	37.17	45.13	38.94	48.67	38.94	48.67
	100	35.40	38.94	29.20	35.40	29.20	35.40
	200	46.02	41.59	30.09	34.51	30.09	34.51
	400	54.87	53.10	41.59	49.56	41.59	49.56
	600	54.87	53.98	43.36	55.75	43.36	55.75
	800	45.13	38.94	33.63	38.94	33.63	38.94
	1000	44.25	39.82	32.74	39.82	32.74	39.82
	1500	43.36	40.71	38.05	40.71	38.05	40.71
0.001	10	5.31	0.00	6.19	0.00	7.96	2.65
	20	24.78	20.35	26.55	18.58	30.09	18.58
	50	38.94	43.36	41.59	43.36	41.59	43.36
	100	53.10	42.48	39.82	44.25	39.82	44.25

	200	47.79	34.51	43.36	35.40	43.36	35.40
	400	51.33	41.59	38.05	38.05	38.05	38.05
	600	51.33	37.17	39.82	38.94	39.82	38.94
	800	41.59	29.20	30.09	30.97	30.09	30.97
	1000	39.82	29.20	30.97	29.20	30.97	29.20
	1500	38.94	28.32	35.40	30.09	35.40	30.09
0.0001	10	5.31	0.00	6.19	0.00	7.96	2.65
	20	5.31	0.00	5.31	0.00	8.85	0.00
	50	38.05	30.09	35.40	28.32	35.40	28.32
	100	41.59	35.40	45.13	43.36	45.13	43.36
	200	40.71	32.74	45.13	38.05	45.13	38.05
	400	42.48	36.28	41.59	36.28	41.59	36.28
	600	43.36	38.94	42.48	42.48	42.48	42.48
	800	43.36	32.74	37.17	32.74	37.17	32.74
	1000	39.82	30.97	33.63	31.86	33.63	31.86
	1500	41.59	33.63	39.82	32.74	39.82	32.74
1	10	50.44	52.21	46.90	49.56	48.67	52.21
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.94	47.79	38.05	47.79	38.05	47.79
	100	41.59	41.59	31.86	38.05	31.86	38.05
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	53.10	53.10	41.59	49.56	41.59	49.56
	600	54.87	53.98	41.59	55.75	41.59	55.75
	800	43.36	46.02	28.32	46.02	28.32	46.02
	1000	43.36	46.90	33.63	46.90	33.63	46.90
	1500	42.48	48.67	37.17	48.67	37.17	48.67
1.5	10	50.44	52.21	46.90	49.56	48.67	52.21
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.94	47.79	38.05	47.79	38.05	47.79
	100	41.59	41.59	31.86	38.05	31.86	38.05
	200	44.25	43.36	31.86	38.05	31.86	38.05
	400	53.10	53.10	41.59	49.56	41.59	49.56
	600	54.87	53.98	41.59	55.75	41.59	55.75
	800	43.36	46.02	28.32	46.02	28.32	46.02
	1000	43.36	46.90	33.63	46.90	33.63	46.90
	1500	42.48	48.67	37.17	48.67	37.17	48.67
5	10	50.44	52.21	46.90	49.56	48.67	52.21
	20	40.71	46.02	39.82	46.02	43.36	46.02
	50	38.94	47.79	38.05	47.79	38.05	47.79

		100	41.59	41.59	31.86	38.05	31.86	38.05
		200	44.25	43.36	31.86	38.05	31.86	38.05
		400	53.10	53.10	41.59	49.56	41.59	49.56
		600	54.87	53.98	41.59	55.75	41.59	55.75
		800	43.36	46.02	28.32	46.02	28.32	46.02
		1000	43.36	46.90	33.63	46.90	33.63	46.90
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	10	10	50.44	52.21	46.90	49.56	48.67	52.21
		20	40.71	46.02	39.82	46.02	43.36	46.02
		50	38.94	47.79	38.05	47.79	38.05	47.79
		100	41.59	41.59	31.86	38.05	31.86	38.05
		200	44.25	43.36	31.86	38.05	31.86	38.05
		400	53.10	53.10	41.59	49.56	41.59	49.56
		600	54.87	53.98	41.59	55.75	41.59	55.75
		800	43.36	46.02	28.32	46.02	28.32	46.02
		1000	43.36	46.90	33.63	46.90	33.63	46.90
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	100	10	50.44	52.21	46.90	49.56	48.67	52.21
		20	40.71	46.02	39.82	46.02	43.36	46.02
		50	38.94	47.79	38.05	47.79	38.05	47.79
		100	41.59	41.59	31.86	38.05	31.86	38.05
		200	44.25	43.36	31.86	38.05	31.86	38.05
		400	53.10	53.10	41.59	49.56	41.59	49.56
		600	54.87	53.98	41.59	55.75	41.59	55.75
		800	43.36	46.02	28.32	46.02	28.32	46.02
		1000	43.36	46.90	33.63	46.90	33.63	46.90
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	1000	10	50.44	52.21	46.90	49.56	48.67	52.21
		20	40.71	46.02	39.82	46.02	43.36	46.02
		50	38.94	47.79	38.05	47.79	38.05	47.79
		100	42.48	41.59	32.74	38.05	32.74	38.05
		200	44.25	43.36	31.86	38.05	31.86	38.05
		400	53.10	53.10	41.59	49.56	41.59	49.56
		600	53.98	53.98	41.59	55.75	41.59	55.75
		800	43.36	46.02	28.32	46.02	28.32	46.02
		1000	42.48	46.90	33.63	46.90	33.63	46.90
		1500	42.48	48.67	37.17	48.67	37.17	48.67

Table 78. Results from the comparison of the measures of dependence versus the MI calculated using the kernel method for detecting the input with the least impact on the output per kernel function, per bandwidth value, and per number of replications.

Bandwidth type	Kernel function	Bandwidth value	Number of replications	Distance correlation		Pearson correlation		R^2_{adj}	
				NIS %	TIS %	NIS %	TIS %	NIS %	TIS %
Silverman	Epanechnikov	NA	10	52.21	53.98	53.10	53.98	54.87	57.52
			20	42.48	46.90	42.48	46.02	43.36	46.02
			50	39.82	49.56	38.94	48.67	38.94	48.67
			100	39.82	40.71	31.86	38.05	32.74	38.05
			200	42.48	38.94	28.32	34.51	28.32	34.51
			400	52.21	52.21	40.71	49.56	40.71	49.56
			600	54.87	52.21	41.59	54.87	41.59	54.87
			800	45.13	46.90	33.63	47.79	33.63	47.79
			1000	41.59	45.13	34.51	46.02	34.51	46.02
			1500	42.48	46.90	37.17	46.90	37.17	46.90
	Normal	NA	10	47.79	50.44	50.44	50.44	52.21	52.21
			20	38.94	45.13	38.05	46.02	38.94	46.02
			50	38.94	46.02	36.28	45.13	36.28	45.13
			100	38.05	38.05	30.09	35.40	30.97	35.40
			200	46.02	42.48	31.86	38.05	31.86	38.05
			400	54.87	53.98	43.36	51.33	43.36	51.33
			600	54.87	53.10	41.59	55.75	41.59	55.75
			800	45.13	47.79	33.63	48.67	33.63	48.67
			1000	41.59	46.02	34.51	46.90	34.51	46.90
			1500	42.48	46.90	37.17	46.90	37.17	46.90
Trial	Epanechnikov	0.1	10	38.05	41.59	42.48	38.94	46.02	39.82
			20	38.94	45.13	43.36	47.79	45.13	47.79
			50	36.28	46.02	35.40	45.13	35.40	45.13
			100	38.05	37.17	30.09	34.51	30.97	34.51
			200	41.59	38.94	29.20	34.51	29.20	34.51
			400	53.10	53.10	41.59	50.44	41.59	50.44
			600	54.87	53.10	41.59	55.75	41.59	55.75
			800	46.02	46.90	32.74	47.79	32.74	47.79
			1000	42.48	45.13	33.63	46.02	33.63	46.02
			1500	40.71	46.02	35.40	46.02	35.40	46.02
		0.2	10	38.94	45.13	46.02	42.48	47.79	43.36
			20	39.82	44.25	39.82	43.36	40.71	43.36
			50	39.82	49.56	38.94	48.67	38.94	48.67

	100	40.71	39.82	30.97	37.17	31.86	37.17
	200	42.48	40.71	30.09	36.28	30.09	36.28
	400	52.21	52.21	40.71	49.56	40.71	49.56
	600	53.98	52.21	40.71	54.87	40.71	54.87
	800	45.13	46.02	31.86	46.90	31.86	46.90
	1000	42.48	45.13	33.63	46.02	33.63	46.02
	1500	42.48	47.79	37.17	47.79	37.17	47.79
0.5	10	50.44	53.98	51.33	53.98	52.21	57.52
	20	41.59	46.02	42.48	46.02	43.36	46.02
	50	38.94	48.67	38.05	47.79	38.05	47.79
	100	38.94	40.71	30.09	38.05	30.97	38.05
	200	43.36	41.59	30.97	37.17	30.97	37.17
	400	52.21	52.21	40.71	49.56	40.71	49.56
	600	53.98	52.21	40.71	54.87	40.71	54.87
	800	40.71	41.59	27.43	40.71	27.43	40.71
	1000	39.82	41.59	29.20	40.71	29.20	40.71
	1500	42.48	47.79	37.17	47.79	37.17	47.79
0.01	10	6.19	0.88	13.27	5.31	8.85	7.08
	20	42.48	37.17	40.71	38.05	43.36	38.94
	50	36.28	44.25	40.71	50.44	40.71	50.44
	100	38.05	39.82	30.09	33.63	31.86	33.63
	200	45.13	39.82	29.20	33.63	29.20	33.63
	400	54.87	53.10	41.59	50.44	41.59	50.44
	600	55.75	53.98	42.48	56.64	42.48	56.64
	800	46.02	36.28	32.74	37.17	32.74	37.17
	1000	43.36	36.28	32.74	37.17	32.74	37.17
	1500	42.48	35.40	37.17	35.40	37.17	35.40
0.001	10	5.31	0.00	13.27	4.42	13.27	8.85
	20	5.31	1.77	5.31	1.77	7.96	3.54
	50	39.82	45.13	38.94	42.48	38.94	42.48
	100	51.33	48.67	39.82	51.33	41.59	51.33
	200	46.02	31.86	43.36	38.05	43.36	38.05
	400	50.44	38.94	37.17	36.28	37.17	36.28
	600	53.10	35.40	38.94	38.05	38.94	38.05
	800	44.25	29.20	30.97	29.20	30.97	29.20
	1000	39.82	27.43	29.20	28.32	29.20	28.32
	1500	38.94	27.43	33.63	29.20	33.63	29.20
0.0001	10	5.31	0.00	13.27	4.42	13.27	8.85
	20	5.31	0.00	7.08	0.88	9.73	3.54

		50	11.50	11.50	9.73	11.50	9.73	11.50
		100	39.82	36.28	44.25	46.90	44.25	46.90
		200	43.36	34.51	45.13	39.82	45.13	39.82
		400	43.36	42.48	44.25	38.94	44.25	38.94
		600	43.36	42.48	44.25	44.25	44.25	44.25
		800	42.48	35.40	36.28	35.40	36.28	35.40
		1000	38.05	34.51	35.40	33.63	35.40	33.63
		1500	40.71	33.63	38.05	33.63	38.05	33.63
	1	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	42.48	46.02	43.36	46.02
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	40.71	39.82	30.97	37.17	31.86	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	41.59	42.48	28.32	43.36	28.32	43.36
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	1.5	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	42.48	46.02	43.36	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	40.71	39.82	30.97	37.17	31.86	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	41.59	42.48	28.32	43.36	28.32	43.36
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	5	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	41.59	46.02	42.48	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	40.71	39.82	30.97	37.17	31.86	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	42.48	43.36	27.43	44.25	27.43	44.25
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	10	10	50.44	53.98	51.33	53.98	53.10	57.52

		20	41.59	46.02	41.59	46.02	42.48	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	40.71	39.82	30.97	37.17	31.86	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	42.48	43.36	27.43	44.25	27.43	44.25
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	100	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	41.59	46.02	42.48	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	40.71	39.82	30.97	37.17	31.86	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	42.48	43.36	27.43	44.25	27.43	44.25
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	1000	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	41.59	46.02	42.48	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	45.13	42.48	35.40	39.82	36.28	39.82
		200	46.02	38.94	32.74	34.51	32.74	34.51
		400	55.75	52.21	44.25	49.56	44.25	49.56
		600	55.75	51.33	43.36	53.98	43.36	53.98
		800	40.71	41.59	27.43	42.48	27.43	42.48
		1000	40.71	44.25	31.86	45.13	31.86	45.13
		1500	42.48	46.90	37.17	46.90	37.17	46.90
Normal	0.1	10	33.63	35.40	38.94	34.51	39.82	35.40
		20	39.82	40.71	40.71	41.59	42.48	41.59
		50	38.94	46.90	32.74	42.48	32.74	42.48
		100	38.94	38.05	29.20	35.40	30.09	35.40
		200	44.25	41.59	31.86	37.17	31.86	37.17
		400	53.10	53.10	41.59	50.44	41.59	50.44
		600	54.87	53.10	41.59	55.75	41.59	55.75
		800	46.02	46.90	32.74	47.79	32.74	47.79
		1000	42.48	46.02	33.63	46.90	33.63	46.90
		1500	40.71	46.02	35.40	46.02	35.40	46.02

	0.2	10	37.17	45.13	45.13	42.48	46.90	43.36
		20	39.82	45.13	38.94	44.25	39.82	44.25
		50	36.28	46.02	35.40	45.13	35.40	45.13
		100	39.82	39.82	30.09	37.17	30.97	37.17
		200	39.82	38.05	27.43	33.63	27.43	33.63
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	54.87	53.10	41.59	55.75	41.59	55.75
		800	46.02	46.90	32.74	47.79	32.74	47.79
		1000	42.48	45.13	33.63	46.02	33.63	46.02
		1500	42.48	46.02	37.17	46.02	37.17	46.02
	0.5	10	49.56	52.21	50.44	52.21	51.33	53.98
		20	41.59	46.02	42.48	45.13	43.36	45.13
		50	39.82	49.56	38.94	48.67	38.94	48.67
		100	39.82	39.82	30.09	37.17	30.97	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	43.36	43.36	28.32	44.25	28.32	44.25
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	0.01	10	49.56	47.79	53.98	47.79	49.56	49.56
		20	40.71	36.28	37.17	37.17	39.82	38.05
		50	36.28	44.25	38.94	48.67	38.94	48.67
		100	35.40	39.82	29.20	35.40	30.97	35.40
		200	45.13	39.82	29.20	33.63	29.20	33.63
		400	54.87	53.10	41.59	50.44	41.59	50.44
		600	54.87	53.10	43.36	55.75	43.36	55.75
		800	45.13	38.05	33.63	38.94	33.63	38.94
		1000	43.36	38.94	32.74	39.82	32.74	39.82
		1500	43.36	39.82	38.05	39.82	38.05	39.82
	0.001	10	5.31	0.00	13.27	3.54	12.39	7.08
		20	23.01	22.12	24.78	21.24	27.43	23.01
		50	38.94	44.25	41.59	43.36	41.59	43.36
		100	53.10	45.13	38.05	46.02	38.94	46.02
		200	46.02	32.74	43.36	35.40	43.36	35.40
		400	50.44	40.71	37.17	37.17	37.17	37.17
		600	51.33	37.17	39.82	38.94	39.82	38.94
		800	42.48	28.32	30.09	30.09	30.09	30.09
		1000	39.82	28.32	30.97	29.20	30.97	29.20

	1500	38.94	28.32	35.40	30.09	35.40	30.09
0.0001	10	5.31	0.00	13.27	4.42	13.27	8.85
	20	5.31	0.00	7.08	0.88	9.73	3.54
	50	37.17	30.97	35.40	29.20	35.40	29.20
	100	41.59	33.63	44.25	42.48	44.25	42.48
	200	39.82	32.74	43.36	38.05	43.36	38.05
	400	44.25	38.94	43.36	37.17	43.36	37.17
	600	45.13	40.71	44.25	42.48	44.25	42.48
	800	44.25	36.28	38.05	36.28	38.05	36.28
	1000	41.59	34.51	34.51	35.40	34.51	35.40
	1500	38.94	32.74	38.05	32.74	38.05	32.74
1	10	50.44	53.98	51.33	53.98	53.10	57.52
	20	41.59	46.02	42.48	46.02	43.36	46.02
	50	39.82	49.56	38.94	48.67	38.94	48.67
	100	40.71	39.82	30.97	37.17	31.86	37.17
	200	43.36	41.59	30.97	37.17	30.97	37.17
	400	52.21	52.21	40.71	49.56	40.71	49.56
	600	53.98	52.21	40.71	54.87	40.71	54.87
	800	42.48	43.36	27.43	44.25	27.43	44.25
	1000	41.59	44.25	32.74	45.13	32.74	45.13
	1500	42.48	48.67	37.17	48.67	37.17	48.67
1.5	10	50.44	53.98	51.33	53.98	53.10	57.52
	20	41.59	46.02	42.48	46.02	43.36	46.02
	50	39.82	49.56	38.94	48.67	38.94	48.67
	100	40.71	39.82	30.97	37.17	31.86	37.17
	200	43.36	41.59	30.97	37.17	30.97	37.17
	400	52.21	52.21	40.71	49.56	40.71	49.56
	600	53.98	52.21	40.71	54.87	40.71	54.87
	800	42.48	43.36	27.43	44.25	27.43	44.25
	1000	41.59	44.25	32.74	45.13	32.74	45.13
	1500	42.48	48.67	37.17	48.67	37.17	48.67
5	10	50.44	53.98	51.33	53.98	53.10	57.52
	20	41.59	46.02	41.59	46.02	42.48	46.90
	50	38.94	48.67	38.05	47.79	38.05	47.79
	100	40.71	39.82	30.97	37.17	31.86	37.17
	200	43.36	41.59	30.97	37.17	30.97	37.17
	400	52.21	52.21	40.71	49.56	40.71	49.56
	600	53.98	52.21	40.71	54.87	40.71	54.87
	800	42.48	43.36	27.43	44.25	27.43	44.25

		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	10	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	41.59	46.02	42.48	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	40.71	39.82	30.97	37.17	31.86	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	42.48	43.36	27.43	44.25	27.43	44.25
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	100	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	41.59	46.02	42.48	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	40.71	39.82	30.97	37.17	31.86	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.98	52.21	40.71	54.87	40.71	54.87
		800	42.48	43.36	27.43	44.25	27.43	44.25
		1000	41.59	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67
	1000	10	50.44	53.98	51.33	53.98	53.10	57.52
		20	41.59	46.02	41.59	46.02	42.48	46.90
		50	38.94	48.67	38.05	47.79	38.05	47.79
		100	41.59	39.82	31.86	37.17	32.74	37.17
		200	43.36	41.59	30.97	37.17	30.97	37.17
		400	52.21	52.21	40.71	49.56	40.71	49.56
		600	53.10	52.21	40.71	54.87	40.71	54.87
		800	42.48	43.36	27.43	44.25	27.43	44.25
		1000	40.71	44.25	32.74	45.13	32.74	45.13
		1500	42.48	48.67	37.17	48.67	37.17	48.67

Table 79. Results from the comparison of the measures of dependence versus the MI calculated using the KNN method for detecting the input with the greatest impact on the output per number of k-nearest neighbors and per number of replications.

Number k-nearest neighbors	Number of replications	Distance correlation	Pearson correlation	R^2_{adj}

		NIS %	TIS %	NIS %	TIS %	NIS %	TIS %
1	10	39.82	43.36	39.82	39.82	39.82	38.94
	20	39.82	35.40	38.05	35.40	38.94	35.40
	50	41.59	36.28	46.02	38.05	46.02	38.05
	100	53.10	40.71	45.13	37.17	45.13	37.17
	200	48.67	45.13	53.98	45.13	53.98	45.13
	400	47.79	42.48	39.82	35.40	39.82	35.40
	600	43.36	47.79	38.94	40.71	38.94	40.71
	800	41.59	49.56	45.13	45.13	45.13	45.13
	1000	46.90	61.06	44.25	55.75	44.25	55.75
	1500	47.79	58.41	46.02	58.41	46.02	58.41
2	10	38.94	40.71	36.28	38.94	36.28	38.05
	20	46.02	42.48	42.48	43.36	43.36	43.36
	50	37.17	44.25	40.71	47.79	40.71	47.79
	100	41.59	33.63	31.86	30.09	31.86	30.09
	200	46.02	35.40	40.71	33.63	40.71	33.63
	400	55.75	44.25	37.17	38.94	37.17	38.94
	600	43.36	35.40	38.05	34.51	38.05	34.51
	800	42.48	37.17	38.94	36.28	38.94	36.28
	1000	43.36	34.51	35.40	30.97	35.40	30.97
	1500	51.33	38.05	46.90	34.51	46.90	34.51
3	10	48.67	49.56	46.02	44.25	46.02	43.36
	20	38.05	45.13	43.36	46.02	44.25	46.02
	50	43.36	44.25	44.25	42.48	44.25	42.48
	100	44.25	39.82	39.82	38.05	39.82	38.05
	200	46.90	35.40	36.28	35.40	36.28	35.40
	400	53.98	33.63	42.48	31.86	42.48	31.86
	600	54.87	28.32	46.90	31.86	46.90	31.86
	800	42.48	26.55	29.20	26.55	29.20	26.55
	1000	36.28	26.55	30.97	25.66	30.97	25.66
	1500	39.82	28.32	36.28	29.20	36.28	29.20
4	10	46.02	47.79	42.48	44.25	42.48	43.36
	20	40.71	38.05	40.71	37.17	41.59	37.17
	50	45.13	46.90	44.25	50.44	44.25	50.44
	100	51.33	45.13	43.36	39.82	43.36	39.82
	200	46.90	38.05	38.05	34.51	38.05	34.51
	400	57.52	39.82	44.25	36.28	44.25	36.28
	600	52.21	34.51	46.02	39.82	46.02	39.82
	800	42.48	32.74	30.97	32.74	30.97	32.74

	1000	42.48	27.43	34.51	25.66	34.51	25.66
	1500	41.59	30.97	36.28	28.32	36.28	28.32
5	10	46.90	49.56	50.44	49.56	50.44	48.67
	20	38.05	38.05	38.05	39.82	38.94	39.82
	50	39.82	46.90	42.48	48.67	42.48	48.67
	100	45.13	39.82	37.17	34.51	37.17	34.51
	200	53.10	39.82	40.71	36.28	40.71	36.28
	400	58.41	43.36	46.90	39.82	46.90	39.82
	600	50.44	36.28	40.71	38.05	40.71	38.05
	800	46.02	33.63	30.97	31.86	30.97	31.86
	1000	39.82	28.32	30.09	28.32	30.09	28.32
	1500	40.71	28.32	31.86	27.43	31.86	27.43
6	10	46.02	51.33	46.90	46.02	46.90	45.13
	20	43.36	42.48	41.59	40.71	42.48	40.71
	50	35.40	46.02	38.94	49.56	38.94	49.56
	100	42.48	44.25	34.51	37.17	34.51	37.17
	200	43.36	36.28	32.74	30.97	32.74	30.97
	400	58.41	46.90	45.13	43.36	45.13	43.36
	600	56.64	43.36	46.90	46.90	46.90	46.90
	800	45.13	30.97	31.86	30.97	31.86	30.97
	1000	41.59	27.43	29.20	27.43	29.20	27.43
	1500	43.36	30.09	36.28	31.86	36.28	31.86
7	10	53.10	53.10	51.33	51.33	51.33	50.44
	20	46.02	46.02	46.02	47.79	46.90	47.79
	50	40.71	47.79	40.71	47.79	40.71	47.79
	100	44.25	43.36	38.05	39.82	38.05	39.82
	200	49.56	42.48	37.17	37.17	37.17	37.17
	400	56.64	46.90	45.13	45.13	45.13	45.13
	600	58.41	39.82	46.90	41.59	46.90	41.59
	800	44.25	34.51	34.51	32.74	34.51	32.74
	1000	38.94	30.09	28.32	28.32	28.32	28.32
	1500	40.71	27.43	35.40	29.20	35.40	29.20
8	10	55.75	53.98	53.98	52.21	52.21	51.33
	20	43.36	45.13	41.59	50.44	42.48	50.44
	50	42.48	40.71	45.13	42.48	45.13	42.48
	100	43.36	42.48	35.40	37.17	35.40	37.17
	200	47.79	40.71	37.17	33.63	37.17	33.63
	400	57.52	48.67	46.02	46.90	46.02	46.90
	600	55.75	50.44	44.25	53.98	44.25	53.98

	800	44.25	34.51	30.97	32.74	30.97	32.74
	1000	41.59	30.97	29.20	30.97	29.20	30.97
	1500	38.05	30.09	34.51	31.86	34.51	31.86
9	10	46.90	53.10	45.13	51.33	48.67	51.33
	20	41.59	53.10	39.82	51.33	40.71	51.33
	50	38.94	41.59	40.71	38.05	40.71	38.05
	100	46.02	43.36	32.74	36.28	32.74	36.28
	200	47.79	38.05	35.40	30.97	35.40	30.97
	400	57.52	50.44	47.79	48.67	47.79	48.67
	600	59.29	50.44	46.02	52.21	46.02	52.21
	800	43.36	35.40	30.09	31.86	30.09	31.86
	1000	40.71	30.09	31.86	30.09	31.86	30.09
	1500	40.71	30.09	35.40	31.86	35.40	31.86

Table 80. Results from the comparison of the measures of dependence versus the MI calculated using the KNN method for detecting the input with the least impact on the output per number of k-nearest neighbors and per number of replications.

Number k-nearest neighbors	Number of replications	Distance correlation		Pearson correlation		R^2_{adj}	
		NIS %	TIS %	NIS %	TIS %	NIS %	TIS %
1	10	46.02	49.56	38.05	41.59	37.17	37.17
	20	45.13	43.36	38.05	39.82	37.17	36.28
	50	49.56	44.25	53.98	45.13	51.33	45.13
	100	58.41	47.79	48.67	41.59	46.90	41.59
	200	52.21	53.10	55.75	50.44	55.75	50.44
	400	50.44	46.02	44.25	39.82	44.25	39.82
	600	44.25	51.33	40.71	44.25	40.71	44.25
	800	46.02	59.29	51.33	55.75	51.33	55.75
	1000	48.67	63.72	47.79	59.29	47.79	59.29
	1500	49.56	61.95	48.67	61.95	48.67	61.95
2	10	43.36	44.25	33.63	38.05	34.51	33.63
	20	50.44	48.67	40.71	46.90	39.82	43.36
	50	43.36	53.10	47.79	57.52	45.13	57.52
	100	46.02	40.71	36.28	36.28	34.51	36.28
	200	54.87	45.13	49.56	42.48	49.56	42.48
	400	64.60	49.56	45.13	45.13	45.13	45.13
	600	49.56	45.13	43.36	44.25	43.36	44.25
	800	47.79	44.25	45.13	43.36	45.13	43.36

	1000	47.79	40.71	40.71	36.28	40.71	36.28
	1500	53.10	45.13	47.79	41.59	47.79	41.59
3	10	47.79	53.98	38.94	44.25	39.82	39.82
	20	45.13	52.21	44.25	50.44	43.36	46.90
	50	47.79	50.44	46.90	47.79	44.25	47.79
	100	46.90	47.79	42.48	45.13	40.71	45.13
	200	51.33	40.71	40.71	41.59	40.71	41.59
	400	60.18	43.36	47.79	41.59	47.79	41.59
	600	61.95	37.17	53.10	40.71	53.10	40.71
	800	46.90	33.63	33.63	34.51	33.63	34.51
	1000	41.59	33.63	36.28	32.74	36.28	32.74
	1500	43.36	33.63	39.82	33.63	39.82	33.63
4	10	48.67	53.98	38.05	47.79	38.94	43.36
	20	46.90	44.25	40.71	40.71	39.82	37.17
	50	50.44	53.10	49.56	55.75	46.90	55.75
	100	52.21	48.67	45.13	44.25	43.36	44.25
	200	49.56	42.48	41.59	39.82	41.59	39.82
	400	60.18	46.90	46.90	43.36	46.90	43.36
	600	57.52	45.13	51.33	50.44	51.33	50.44
	800	47.79	39.82	35.40	39.82	35.40	39.82
	1000	48.67	34.51	39.82	33.63	39.82	33.63
	1500	44.25	37.17	38.94	35.40	38.94	35.40
5	10	48.67	53.10	46.02	50.44	46.90	46.02
	20	47.79	48.67	41.59	46.90	40.71	43.36
	50	47.79	55.75	48.67	56.64	46.02	56.64
	100	47.79	46.02	39.82	40.71	38.05	40.71
	200	54.87	45.13	43.36	42.48	43.36	42.48
	400	61.95	50.44	50.44	46.90	50.44	46.90
	600	56.64	46.90	46.90	48.67	46.90	48.67
	800	51.33	39.82	35.40	38.05	35.40	38.05
	1000	47.79	36.28	37.17	37.17	37.17	37.17
	1500	45.13	35.40	36.28	35.40	36.28	35.40
6	10	49.56	57.52	42.48	47.79	42.48	43.36
	20	47.79	46.90	40.71	43.36	39.82	39.82
	50	43.36	53.98	47.79	58.41	45.13	58.41
	100	47.79	53.10	39.82	45.13	38.05	45.13
	200	48.67	44.25	38.94	38.05	38.94	38.05
	400	61.06	54.87	47.79	51.33	47.79	51.33
	600	62.83	51.33	53.10	54.87	53.10	54.87

	800	51.33	38.94	37.17	38.94	37.17	38.94
	1000	48.67	35.40	36.28	36.28	36.28	36.28
	1500	50.44	38.05	43.36	39.82	43.36	39.82
7	10	54.87	57.52	46.90	53.10	46.90	48.67
	20	50.44	53.10	45.13	53.10	44.25	49.56
	50	47.79	54.87	48.67	55.75	46.02	55.75
	100	46.90	51.33	40.71	46.90	38.94	46.90
	200	53.10	47.79	41.59	43.36	41.59	43.36
	400	58.41	53.98	46.90	52.21	46.90	52.21
	600	61.06	46.90	50.44	49.56	50.44	49.56
	800	51.33	42.48	41.59	40.71	41.59	40.71
	1000	46.90	38.05	36.28	37.17	36.28	37.17
	1500	47.79	37.17	42.48	38.94	42.48	38.94
8	10	60.18	60.18	50.44	55.75	51.33	51.33
	20	53.10	55.75	46.02	57.52	45.13	53.98
	50	50.44	49.56	53.98	51.33	51.33	51.33
	100	45.13	48.67	37.17	44.25	35.40	44.25
	200	51.33	46.90	41.59	40.71	41.59	40.71
	400	61.95	55.75	50.44	53.98	50.44	53.98
	600	61.95	57.52	50.44	61.95	50.44	61.95
	800	50.44	41.59	37.17	39.82	37.17	39.82
	1000	48.67	39.82	37.17	40.71	37.17	40.71
	1500	45.13	38.05	41.59	39.82	41.59	39.82
9	10	46.90	56.64	43.36	53.10	44.25	48.67
	20	47.79	61.06	40.71	57.52	39.82	53.98
	50	48.67	51.33	50.44	47.79	47.79	47.79
	100	50.44	53.10	37.17	45.13	35.40	45.13
	200	51.33	44.25	39.82	38.05	39.82	38.05
	400	61.06	58.41	51.33	56.64	51.33	56.64
	600	65.49	58.41	52.21	61.06	52.21	61.06
	800	50.44	43.36	37.17	39.82	37.17	39.82
	1000	46.02	37.17	38.05	38.05	38.05	38.05
	1500	46.90	37.17	41.59	38.94	41.59	38.94

Table 81. Results from the comparison of the measures of dependence versus the MI calculated using the fuzzy-histogram based method for detecting the input with the greatest impact on the output per membership function, per number of fuzzy subsets, and per number of replications.

Membership function	Number of fuzzy subsets	Number of replications	Distance correlation		Pearson correlation		R^2_{adj}	
			NIS %	TIS %	NIS %	TIS %	NIS %	TIS %
Cosine	2	10	26.55	17.70	10.62	7.08	14.16	7.08
		20	23.89	15.04	18.58	8.85	19.47	8.85
		50	29.20	13.27	21.24	7.96	21.24	7.96
		100	36.28	21.24	19.47	10.62	19.47	10.62
		200	37.17	21.24	19.47	8.85	19.47	8.85
		400	44.25	23.89	30.97	13.27	30.97	13.27
		600	41.59	24.78	28.32	12.39	28.32	12.39
		800	42.48	28.32	25.66	17.70	25.66	17.70
		1000	43.36	29.20	25.66	18.58	25.66	18.58
		1500	38.05	29.20	25.66	18.58	25.66	18.58
	5	10	30.09	19.47	14.16	12.39	17.70	12.39
		20	23.01	12.39	19.47	7.96	20.35	7.96
		50	33.63	12.39	27.43	8.85	27.43	8.85
		100	33.63	20.35	16.81	9.73	16.81	9.73
		200	38.05	21.24	22.12	10.62	22.12	10.62
		400	43.36	24.78	30.09	12.39	30.09	12.39
		600	39.82	24.78	28.32	12.39	28.32	12.39
		800	42.48	29.20	27.43	18.58	27.43	18.58
		1000	45.13	29.20	27.43	18.58	27.43	18.58
		1500	39.82	29.20	25.66	18.58	25.66	18.58
	10	10	30.97	20.35	15.04	13.27	18.58	13.27
		20	22.12	15.04	18.58	8.85	19.47	8.85
		50	35.40	13.27	27.43	7.96	27.43	7.96
		100	36.28	23.01	19.47	10.62	19.47	10.62
		200	43.36	21.24	27.43	10.62	27.43	10.62
		400	45.13	24.78	31.86	12.39	31.86	12.39
		600	38.94	25.66	29.20	13.27	29.20	13.27
		800	43.36	28.32	26.55	17.70	26.55	17.70
		1000	45.13	28.32	25.66	17.70	25.66	17.70
		1500	42.48	29.20	28.32	18.58	28.32	18.58
	25	10	23.01	22.12	9.73	15.04	10.62	15.04
		20	29.20	19.47	27.43	13.27	24.78	13.27
		50	33.63	13.27	29.20	9.73	29.20	9.73
		100	39.82	20.35	24.78	13.27	24.78	13.27
		200	42.48	22.12	30.09	11.50	30.09	11.50
		400	46.90	23.89	38.94	13.27	38.94	13.27

		600	41.59	24.78	35.40	14.16	35.40	14.16
		800	46.02	28.32	30.97	17.70	30.97	17.70
		1000	47.79	28.32	30.09	17.70	30.09	17.70
		1500	44.25	29.20	33.63	20.35	33.63	20.35
	50	10	26.55	15.93	11.50	7.08	14.16	7.08
		20	30.09	18.58	30.09	14.16	26.55	14.16
		50	38.94	18.58	30.97	15.04	30.97	15.04
		100	36.28	19.47	30.97	18.58	30.97	18.58
		200	43.36	22.12	34.51	22.12	34.51	22.12
		400	49.56	21.24	43.36	14.16	43.36	14.16
		600	43.36	22.12	37.17	13.27	37.17	13.27
		800	41.59	29.20	40.71	18.58	40.71	18.58
		1000	45.13	30.97	41.59	20.35	41.59	20.35
		1500	46.90	31.86	45.13	23.01	45.13	23.01
	100	10	23.01	17.70	12.39	10.62	14.16	10.62
		20	38.05	22.12	34.51	17.70	30.09	17.70
		50	41.59	20.35	35.40	18.58	35.40	18.58
		100	36.28	25.66	38.05	24.78	38.05	24.78
		200	43.36	27.43	44.25	27.43	44.25	27.43
		400	44.25	24.78	43.36	23.01	43.36	23.01
		600	43.36	23.89	40.71	20.35	40.71	20.35
		800	38.94	30.09	42.48	23.89	42.48	23.89
		1000	38.94	30.97	43.36	24.78	43.36	24.78
		1500	42.48	31.86	40.71	26.55	40.71	26.55
	200	10	35.40	25.66	28.32	18.58	30.09	17.70
		20	35.40	20.35	37.17	17.70	34.51	17.70
		50	44.25	20.35	39.82	16.81	39.82	16.81
		100	40.71	26.55	42.48	27.43	42.48	27.43
		200	43.36	31.86	42.48	30.09	42.48	30.09
		400	44.25	29.20	47.79	29.20	47.79	29.20
		600	40.71	28.32	44.25	23.89	44.25	23.89
		800	41.59	34.51	48.67	28.32	48.67	28.32
		1000	45.13	34.51	49.56	29.20	49.56	29.20
		1500	46.02	32.74	49.56	28.32	49.56	28.32
	500	10	42.48	29.20	37.17	18.58	38.05	18.58
		20	39.82	28.32	39.82	23.01	36.28	23.01
		50	51.33	23.01	47.79	19.47	47.79	19.47
		100	38.05	30.09	45.13	32.74	45.13	32.74
		200	46.90	30.09	43.36	30.09	43.36	30.09

		400	43.36	30.97	47.79	30.97	47.79	30.97
		600	41.59	29.20	49.56	27.43	49.56	27.43
		800	46.02	33.63	49.56	27.43	49.56	27.43
		1000	49.56	34.51	52.21	30.09	52.21	30.09
		1500	48.67	34.51	52.21	31.86	52.21	31.86
	1000	10	41.59	30.97	37.17	23.89	36.28	23.01
		20	38.05	30.09	39.82	27.43	35.40	27.43
		50	46.90	23.89	38.05	18.58	38.05	18.58
		100	46.90	34.51	48.67	35.40	48.67	35.40
		200	46.02	31.86	35.40	30.09	35.40	30.09
		400	41.59	31.86	43.36	29.20	43.36	29.20
		600	39.82	30.09	46.90	29.20	46.90	29.20
		800	44.25	33.63	50.44	30.09	50.44	30.09
		1000	45.13	38.05	56.64	37.17	56.64	37.17
		1500	45.13	35.40	54.87	32.74	54.87	32.74
fCrisp	2	10	11.50	2.65	6.19	1.77	9.73	1.77
		20	30.09	14.16	23.89	10.62	23.89	10.62
		50	31.86	17.70	23.89	12.39	23.89	12.39
		100	45.13	23.01	30.97	14.16	30.97	14.16
		200	46.90	22.12	30.09	12.39	30.09	12.39
		400	46.90	22.12	33.63	14.16	33.63	14.16
		600	46.90	23.89	31.86	12.39	31.86	12.39
		800	51.33	30.09	32.74	19.47	32.74	19.47
		1000	49.56	30.97	30.97	22.12	30.97	22.12
		1500	45.13	30.09	30.09	21.24	30.09	21.24
	5	10	38.94	17.70	28.32	16.81	31.86	16.81
		20	30.09	13.27	28.32	11.50	29.20	11.50
		50	36.28	14.16	30.09	12.39	30.09	12.39
		100	38.94	20.35	19.47	7.96	19.47	7.96
		200	45.13	21.24	25.66	11.50	25.66	11.50
		400	46.02	25.66	33.63	14.16	33.63	14.16
		600	46.90	25.66	34.51	15.04	34.51	15.04
		800	45.13	31.86	30.09	21.24	30.09	21.24
		1000	47.79	31.86	29.20	21.24	29.20	21.24
		1500	42.48	30.97	23.89	22.12	23.89	22.12
	10	10	32.74	15.93	23.01	13.27	25.66	12.39
		20	32.74	17.70	28.32	12.39	27.43	12.39
		50	38.05	16.81	31.86	9.73	31.86	9.73
		100	41.59	23.89	23.89	13.27	23.89	13.27

		200	49.56	21.24	30.97	11.50	30.97	11.50
		400	47.79	25.66	34.51	15.93	34.51	15.93
		600	46.02	24.78	36.28	13.27	36.28	13.27
		800	46.90	30.09	31.86	17.70	31.86	17.70
		1000	46.02	29.20	26.55	16.81	26.55	16.81
		1500	46.02	30.09	28.32	19.47	28.32	19.47
	25	10	28.32	15.04	20.35	11.50	23.01	11.50
		20	36.28	19.47	33.63	17.70	34.51	17.70
		50	37.17	17.70	31.86	14.16	31.86	14.16
		100	37.17	23.01	23.89	15.04	23.89	15.04
		200	41.59	20.35	33.63	15.93	33.63	15.93
		400	47.79	21.24	39.82	15.93	39.82	15.93
		600	45.13	22.12	35.40	13.27	35.40	13.27
		800	43.36	30.97	33.63	22.12	33.63	22.12
		1000	44.25	31.86	35.40	19.47	35.40	19.47
		1500	44.25	30.97	38.94	22.12	38.94	22.12
	50	10	30.09	16.81	18.58	9.73	19.47	9.73
		20	32.74	19.47	34.51	15.04	30.97	15.04
		50	42.48	17.70	33.63	14.16	33.63	14.16
		100	39.82	23.01	38.05	23.89	38.05	23.89
		200	40.71	24.78	34.51	23.01	34.51	23.01
		400	48.67	23.01	46.02	15.93	46.02	15.93
		600	45.13	24.78	40.71	15.93	40.71	15.93
		800	41.59	29.20	40.71	22.12	40.71	22.12
		1000	44.25	30.09	40.71	23.01	40.71	23.01
		1500	46.90	33.63	43.36	24.78	43.36	24.78
	100	10	30.09	17.70	17.70	12.39	19.47	11.50
		20	38.94	22.12	37.17	18.58	33.63	18.58
		50	46.90	21.24	46.90	21.24	46.90	21.24
		100	36.28	27.43	41.59	26.55	41.59	26.55
		200	41.59	25.66	38.94	27.43	38.94	27.43
		400	43.36	29.20	44.25	27.43	44.25	27.43
		600	42.48	27.43	45.13	23.89	45.13	23.89
		800	40.71	30.97	42.48	24.78	42.48	24.78
		1000	41.59	31.86	44.25	26.55	44.25	26.55
		1500	41.59	32.74	43.36	28.32	43.36	28.32
	200	10	29.20	13.27	18.58	6.19	20.35	6.19
		20	24.78	17.70	25.66	13.27	23.89	13.27
		50	45.13	22.12	39.82	18.58	39.82	18.58

		100	40.71	28.32	38.94	27.43	38.94	27.43
		200	42.48	30.97	46.90	33.63	46.90	33.63
		400	46.02	28.32	47.79	28.32	47.79	28.32
		600	43.36	30.09	45.13	28.32	45.13	28.32
		800	45.13	34.51	48.67	28.32	48.67	28.32
		1000	46.02	36.28	48.67	30.97	48.67	30.97
		1500	45.13	33.63	50.44	29.20	50.44	29.20
	500	10	20.35	14.16	15.04	8.85	15.93	7.96
		20	22.12	13.27	21.24	11.50	20.35	11.50
		50	49.56	23.01	48.67	22.12	48.67	22.12
		100	42.48	36.28	43.36	32.74	43.36	32.74
		200	46.90	31.86	34.51	26.55	34.51	26.55
		400	42.48	29.20	46.90	30.97	46.90	30.97
		600	38.05	29.20	46.90	26.55	46.90	26.55
		800	46.02	33.63	45.13	28.32	45.13	28.32
		1000	47.79	35.40	52.21	32.74	52.21	32.74
		1500	49.56	35.40	51.33	32.74	51.33	32.74
	1000	10	16.81	12.39	11.50	7.08	12.39	6.19
		20	18.58	13.27	19.47	10.62	18.58	10.62
		50	41.59	30.97	35.40	25.66	35.40	25.66
		100	44.25	38.05	46.90	36.28	46.90	36.28
		200	47.79	33.63	48.67	36.28	48.67	36.28
		400	43.36	30.09	50.44	30.97	50.44	30.97
		600	41.59	30.97	46.90	30.09	46.90	30.09
		800	44.25	34.51	52.21	29.20	52.21	29.20
		1000	45.13	38.05	53.98	36.28	53.98	36.28
		1500	42.48	33.63	46.90	30.97	46.90	30.97
Triangular	2	10	26.55	17.70	10.62	8.85	14.16	8.85
		20	23.01	13.27	17.70	7.08	18.58	7.08
		50	29.20	13.27	23.01	7.96	23.01	7.96
		100	34.51	22.12	15.93	9.73	15.93	9.73
		200	38.05	20.35	18.58	7.96	18.58	7.96
		400	38.05	23.89	26.55	11.50	26.55	11.50
		600	38.94	24.78	25.66	12.39	25.66	12.39
		800	43.36	27.43	26.55	16.81	26.55	16.81
		1000	44.25	28.32	26.55	17.70	26.55	17.70
		1500	38.05	28.32	25.66	17.70	25.66	17.70
	5	10	25.66	18.58	9.73	11.50	13.27	11.50
		20	21.24	14.16	17.70	7.96	18.58	7.96

		50	32.74	13.27	24.78	7.96	24.78	7.96
		100	31.86	21.24	16.81	10.62	16.81	10.62
		200	38.05	21.24	22.12	10.62	22.12	10.62
		400	43.36	23.89	30.09	13.27	30.09	13.27
		600	39.82	24.78	28.32	12.39	28.32	12.39
		800	41.59	28.32	26.55	17.70	26.55	17.70
		1000	44.25	29.20	24.78	18.58	24.78	18.58
		1500	38.94	29.20	24.78	18.58	24.78	18.58
	10	10	23.01	18.58	7.08	11.50	10.62	11.50
		20	23.01	15.93	19.47	9.73	20.35	9.73
		50	37.17	13.27	29.20	7.96	29.20	7.96
		100	36.28	22.12	19.47	11.50	19.47	11.50
		200	42.48	21.24	26.55	10.62	26.55	10.62
		400	44.25	24.78	30.97	12.39	30.97	12.39
		600	38.94	27.43	29.20	15.04	29.20	15.04
		800	42.48	28.32	27.43	17.70	27.43	17.70
		1000	45.13	28.32	25.66	17.70	25.66	17.70
		1500	42.48	29.20	28.32	18.58	28.32	18.58
	25	10	23.01	17.70	9.73	10.62	10.62	10.62
		20	27.43	16.81	23.89	12.39	21.24	12.39
		50	34.51	14.16	28.32	8.85	28.32	8.85
		100	40.71	22.12	25.66	13.27	25.66	13.27
		200	43.36	23.01	30.97	12.39	30.97	12.39
		400	45.13	23.89	38.94	13.27	38.94	13.27
		600	41.59	25.66	35.40	15.04	35.40	15.04
		800	46.02	28.32	30.97	17.70	30.97	17.70
		1000	47.79	28.32	28.32	17.70	28.32	17.70
		1500	46.02	29.20	33.63	20.35	33.63	20.35
	50	10	26.55	17.70	11.50	8.85	14.16	8.85
		20	30.97	18.58	29.20	15.93	25.66	15.93
		50	37.17	15.93	29.20	12.39	29.20	12.39
		100	37.17	21.24	26.55	16.81	26.55	16.81
		200	43.36	23.01	36.28	21.24	36.28	21.24
		400	47.79	21.24	41.59	14.16	41.59	14.16
		600	42.48	23.01	38.05	12.39	38.05	12.39
		800	42.48	30.97	38.05	20.35	38.05	20.35
		1000	45.13	30.09	39.82	17.70	39.82	17.70
		1500	46.02	30.97	42.48	22.12	42.48	22.12
	100	10	25.66	16.81	11.50	7.96	13.27	7.96

		20	34.51	19.47	32.74	16.81	29.20	16.81
		50	43.36	19.47	37.17	17.70	37.17	17.70
		100	38.94	24.78	35.40	22.12	35.40	22.12
		200	42.48	26.55	39.82	26.55	39.82	26.55
		400	45.13	23.89	44.25	19.47	44.25	19.47
		600	43.36	23.89	40.71	17.70	40.71	17.70
		800	40.71	30.97	41.59	23.01	41.59	23.01
		1000	41.59	30.97	43.36	24.78	43.36	24.78
		1500	43.36	30.97	41.59	25.66	41.59	25.66
	200	10	30.97	24.78	23.89	17.70	25.66	16.81
		20	35.40	22.12	37.17	19.47	34.51	19.47
		50	43.36	19.47	38.94	17.70	38.94	17.70
		100	38.94	26.55	38.94	25.66	38.94	25.66
		200	43.36	31.86	40.71	30.09	40.71	30.09
		400	45.13	30.09	46.90	28.32	46.90	28.32
		600	39.82	28.32	43.36	23.89	43.36	23.89
		800	42.48	33.63	47.79	27.43	47.79	27.43
		1000	42.48	34.51	46.90	29.20	46.90	29.20
		1500	44.25	32.74	46.02	28.32	46.02	28.32
	500	10	39.82	25.66	35.40	15.04	36.28	15.04
		20	39.82	26.55	39.82	21.24	36.28	21.24
		50	50.44	20.35	43.36	16.81	43.36	16.81
		100	39.82	29.20	46.90	30.09	46.90	30.09
		200	48.67	30.97	45.13	29.20	45.13	29.20
		400	44.25	30.09	46.90	30.09	46.90	30.09
		600	42.48	28.32	50.44	26.55	50.44	26.55
		800	44.25	32.74	49.56	26.55	49.56	26.55
		1000	50.44	34.51	52.21	30.09	52.21	30.09
		1500	47.79	35.40	51.33	30.97	51.33	30.97
	1000	10	36.28	27.43	33.63	20.35	32.74	19.47
		20	38.94	27.43	38.94	24.78	34.51	24.78
		50	47.79	24.78	42.48	19.47	42.48	19.47
		100	48.67	30.97	52.21	30.09	52.21	30.09
		200	46.02	32.74	37.17	29.20	37.17	29.20
		400	40.71	31.86	42.48	30.09	42.48	30.09
		600	41.59	30.97	46.90	28.32	46.90	28.32
		800	44.25	33.63	50.44	30.09	50.44	30.09
		1000	48.67	36.28	54.87	33.63	54.87	33.63
		1500	44.25	35.40	54.87	32.74	54.87	32.74

Table 82. Results from the comparison of the measures of dependence versus the MI calculated using the fuzzy-histogram based method for detecting the input with the least impact on the output per membership function, per number of fuzzy subsets, and per number of replications.

Membership function	Number of fuzzy subsets	Number of replications	Distance correlation		Pearson correlation		R^2_{adj}	
			NIS %	TIS %	NIS %	NIS %	TIS %	NIS %
Cosine	2	10	30.09	23.89	11.50	11.50	11.50	7.08
		20	31.86	23.89	21.24	15.04	20.35	11.50
		50	37.17	21.24	30.09	15.93	27.43	15.93
		100	40.71	30.09	24.78	20.35	23.01	20.35
		200	42.48	30.09	25.66	17.70	25.66	17.70
		400	47.79	32.74	35.40	23.01	35.40	23.01
		600	46.90	33.63	34.51	22.12	34.51	22.12
		800	47.79	36.28	31.86	26.55	31.86	26.55
		1000	48.67	37.17	31.86	27.43	31.86	27.43
		1500	44.25	37.17	32.74	26.55	32.74	26.55
	5	10	33.63	26.55	15.04	15.93	15.04	11.50
		20	30.97	22.12	22.12	15.04	21.24	11.50
		50	38.05	20.35	32.74	16.81	30.09	16.81
		100	40.71	30.09	24.78	19.47	23.01	19.47
		200	43.36	29.20	28.32	18.58	28.32	18.58
		400	46.90	32.74	34.51	21.24	34.51	21.24
		600	44.25	33.63	33.63	22.12	33.63	22.12
		800	47.79	37.17	33.63	27.43	33.63	27.43
		1000	50.44	37.17	33.63	27.43	33.63	27.43
		1500	46.02	37.17	32.74	26.55	32.74	26.55
	10	10	33.63	26.55	15.93	15.93	15.93	11.50
		20	30.09	23.01	21.24	14.16	20.35	10.62
		50	39.82	22.12	32.74	16.81	30.09	16.81
		100	40.71	31.86	24.78	19.47	23.01	19.47
		200	45.13	29.20	30.09	18.58	30.09	18.58
		400	48.67	33.63	36.28	22.12	36.28	22.12
		600	43.36	34.51	34.51	23.01	34.51	23.01
		800	47.79	36.28	31.86	26.55	31.86	26.55
		1000	49.56	36.28	30.97	26.55	30.97	26.55
		1500	46.02	37.17	32.74	26.55	32.74	26.55

	25	10	28.32	29.20	9.73	18.58	9.73	14.16
		20	35.40	25.66	26.55	16.81	25.66	13.27
		50	38.05	23.01	34.51	19.47	31.86	19.47
		100	41.59	27.43	27.43	21.24	25.66	21.24
		200	43.36	29.20	31.86	18.58	31.86	18.58
		400	50.44	32.74	43.36	23.01	43.36	23.01
		600	46.02	33.63	40.71	23.89	40.71	23.89
		800	50.44	36.28	36.28	26.55	36.28	26.55
		1000	52.21	36.28	35.40	26.55	35.40	26.55
		1500	47.79	37.17	38.05	28.32	38.05	28.32
	50	10	30.97	22.12	12.39	11.50	13.27	7.08
		20	34.51	23.01	27.43	15.93	26.55	12.39
		50	43.36	25.66	36.28	22.12	33.63	22.12
		100	39.82	27.43	34.51	25.66	32.74	25.66
		200	43.36	29.20	35.40	29.20	35.40	29.20
		400	53.10	29.20	47.79	23.01	47.79	23.01
		600	47.79	30.09	42.48	22.12	42.48	22.12
		800	46.02	36.28	46.02	26.55	46.02	26.55
		1000	46.90	37.17	44.25	27.43	44.25	27.43
		1500	46.02	38.05	45.13	29.20	45.13	29.20
	100	10	27.43	24.78	12.39	14.16	12.39	9.73
		20	41.59	25.66	30.97	18.58	30.09	15.04
		50	43.36	24.78	38.05	23.01	35.40	23.01
		100	38.05	30.97	40.71	29.20	38.94	29.20
		200	41.59	31.86	42.48	31.86	42.48	31.86
		400	47.79	31.86	47.79	29.20	47.79	29.20
		600	46.90	30.97	45.13	26.55	45.13	26.55
		800	44.25	36.28	46.90	30.09	46.90	30.09
		1000	41.59	36.28	45.13	30.09	45.13	30.09
		1500	44.25	37.17	42.48	31.86	42.48	31.86
	200	10	38.05	33.63	25.66	22.12	25.66	17.70
		20	40.71	26.55	35.40	21.24	34.51	17.70
		50	44.25	25.66	40.71	22.12	38.05	22.12
		100	41.59	31.86	44.25	32.74	42.48	32.74
		200	42.48	35.40	41.59	33.63	41.59	33.63
		400	48.67	34.51	52.21	33.63	52.21	33.63
		600	46.02	35.40	49.56	30.97	49.56	30.97
		800	45.13	38.94	50.44	32.74	50.44	32.74
		1000	46.02	38.05	48.67	31.86	48.67	31.86

		1500	47.79	38.05	50.44	32.74	50.44	32.74
	500	10	45.13	35.40	35.40	23.01	36.28	18.58
		20	43.36	32.74	37.17	25.66	36.28	22.12
		50	49.56	29.20	46.02	25.66	43.36	25.66
		100	38.94	32.74	46.90	36.28	45.13	36.28
		200	43.36	31.86	40.71	31.86	40.71	31.86
		400	46.90	34.51	50.44	33.63	50.44	33.63
		600	42.48	34.51	49.56	31.86	49.56	31.86
		800	48.67	39.82	52.21	33.63	52.21	33.63
		1000	50.44	40.71	51.33	34.51	51.33	34.51
		1500	49.56	38.05	53.98	34.51	53.98	34.51
	1000	10	49.56	36.28	36.28	25.66	36.28	21.24
		20	40.71	33.63	36.28	28.32	35.40	24.78
		50	46.90	25.66	38.05	20.35	35.40	20.35
		100	48.67	32.74	50.44	34.51	48.67	34.51
		200	46.02	34.51	36.28	32.74	36.28	32.74
		400	46.02	35.40	48.67	32.74	48.67	32.74
		600	43.36	35.40	52.21	34.51	52.21	34.51
		800	48.67	40.71	53.98	36.28	53.98	36.28
		1000	45.13	41.59	54.87	38.94	54.87	38.94
		1500	46.02	39.82	57.52	36.28	57.52	36.28
Crisp	2	10	14.16	8.85	6.19	4.42	6.19	0.00
		20	36.28	20.35	24.78	14.16	23.89	10.62
		50	39.82	23.89	32.74	18.58	30.09	18.58
		100	48.67	30.09	34.51	22.12	32.74	22.12
		200	52.21	31.86	34.51	21.24	34.51	21.24
		400	51.33	30.97	38.94	23.01	38.94	23.01
		600	50.44	32.74	36.28	21.24	36.28	21.24
		800	54.87	38.05	36.28	28.32	36.28	28.32
		1000	53.10	38.05	33.63	30.09	33.63	30.09
		1500	49.56	37.17	34.51	28.32	34.51	28.32
	5	10	40.71	23.01	29.20	19.47	29.20	15.04
		20	38.05	23.89	30.97	18.58	30.09	15.04
		50	40.71	22.12	35.40	20.35	32.74	20.35
		100	45.13	28.32	26.55	16.81	24.78	16.81
		200	50.44	29.20	30.97	19.47	30.97	19.47
		400	49.56	33.63	38.05	23.01	38.05	23.01
		600	52.21	34.51	40.71	24.78	40.71	24.78
		800	50.44	39.82	36.28	30.09	36.28	30.09

		1000	53.98	39.82	36.28	30.09	36.28	30.09
		1500	50.44	38.94	31.86	30.09	31.86	30.09
	10	10	35.40	18.58	22.12	14.16	22.12	9.73
		20	39.82	23.01	28.32	15.93	27.43	12.39
		50	45.13	24.78	39.82	17.70	37.17	17.70
		100	45.13	31.86	28.32	22.12	26.55	22.12
		200	50.44	29.20	32.74	19.47	32.74	19.47
		400	52.21	33.63	39.82	24.78	39.82	24.78
		600	50.44	33.63	41.59	23.01	41.59	23.01
		800	51.33	38.05	37.17	26.55	37.17	26.55
		1000	50.44	37.17	31.86	25.66	31.86	25.66
		1500	49.56	38.05	32.74	27.43	32.74	27.43
	25	10	32.74	23.01	21.24	17.70	21.24	13.27
		20	42.48	25.66	34.51	21.24	33.63	17.70
		50	42.48	26.55	38.05	23.01	35.40	23.01
		100	41.59	30.97	28.32	23.01	26.55	23.01
		200	43.36	27.43	36.28	23.01	36.28	23.01
		400	50.44	30.09	43.36	25.66	43.36	25.66
		600	49.56	30.97	40.71	23.01	40.71	23.01
		800	48.67	38.94	39.82	30.97	39.82	30.97
		1000	47.79	39.82	39.82	28.32	39.82	28.32
		1500	46.02	38.94	41.59	30.09	41.59	30.09
	50	10	32.74	26.55	16.81	15.93	16.81	11.50
		20	34.51	23.01	29.20	15.93	28.32	12.39
		50	47.79	24.78	39.82	21.24	37.17	21.24
		100	43.36	30.97	42.48	31.86	40.71	31.86
		200	42.48	30.97	36.28	29.20	36.28	29.20
		400	51.33	30.09	49.56	23.89	49.56	23.89
		600	49.56	31.86	46.02	23.89	46.02	23.89
		800	46.02	35.40	46.02	29.20	46.02	29.20
		1000	45.13	35.40	42.48	29.20	42.48	29.20
		1500	46.90	38.94	44.25	30.09	44.25	30.09
	100	10	32.74	27.43	15.93	18.58	16.81	14.16
		20	43.36	30.09	34.51	23.01	33.63	19.47
		50	44.25	25.66	45.13	25.66	42.48	25.66
		100	38.05	33.63	44.25	31.86	42.48	31.86
		200	41.59	29.20	38.94	30.97	38.94	30.97
		400	47.79	34.51	49.56	31.86	49.56	31.86
		600	46.02	33.63	49.56	29.20	49.56	29.20

		800	46.02	36.28	46.02	30.09	46.02	30.09
		1000	45.13	36.28	46.02	30.09	46.02	30.09
		1500	45.13	36.28	46.02	30.97	46.02	30.97
	200	10	33.63	24.78	18.58	15.93	18.58	11.50
		20	30.97	27.43	24.78	19.47	23.89	15.93
		50	46.90	28.32	41.59	24.78	38.94	24.78
		100	42.48	35.40	41.59	34.51	39.82	34.51
		200	42.48	34.51	47.79	36.28	47.79	36.28
		400	50.44	34.51	52.21	33.63	52.21	33.63
		600	47.79	34.51	49.56	31.86	49.56	31.86
		800	48.67	38.05	50.44	31.86	50.44	31.86
		1000	47.79	38.94	50.44	32.74	50.44	32.74
		1500	47.79	38.05	52.21	32.74	52.21	32.74
	500	10	23.89	26.55	14.16	18.58	14.16	14.16
		20	24.78	24.78	18.58	19.47	17.70	15.93
		50	52.21	27.43	51.33	27.43	48.67	27.43
		100	43.36	36.28	45.13	33.63	43.36	33.63
		200	44.25	33.63	32.74	28.32	32.74	28.32
		400	45.13	34.51	48.67	35.40	48.67	35.40
		600	40.71	33.63	51.33	30.97	51.33	30.97
		800	50.44	38.94	48.67	32.74	48.67	32.74
		1000	48.67	41.59	51.33	37.17	51.33	37.17
		1500	51.33	38.94	53.98	35.40	53.98	35.40
	1000	10	19.47	19.47	9.73	10.62	9.73	6.19
		20	21.24	21.24	16.81	15.93	15.93	12.39
		50	43.36	36.28	38.05	30.09	35.40	30.09
		100	46.02	38.94	49.56	38.94	47.79	38.94
		200	48.67	38.94	49.56	40.71	49.56	40.71
		400	44.25	32.74	52.21	33.63	52.21	33.63
		600	46.02	38.05	52.21	37.17	52.21	37.17
		800	48.67	38.94	55.75	32.74	55.75	32.74
		1000	46.02	39.82	53.98	37.17	53.98	37.17
		1500	44.25	37.17	50.44	33.63	50.44	33.63
Triangular	2	10	29.20	23.89	11.50	13.27	11.50	8.85
		20	30.97	23.01	20.35	14.16	19.47	10.62
		50	36.28	21.24	30.97	15.93	28.32	15.93
		100	39.82	30.09	22.12	18.58	20.35	18.58
		200	44.25	29.20	25.66	16.81	25.66	16.81
		400	43.36	32.74	32.74	21.24	32.74	21.24

		600	44.25	33.63	31.86	22.12	31.86	22.12
		800	48.67	35.40	32.74	25.66	32.74	25.66
		1000	49.56	36.28	32.74	26.55	32.74	26.55
		1500	44.25	36.28	32.74	25.66	32.74	25.66
	5	10	28.32	25.66	10.62	15.04	10.62	10.62
		20	29.20	23.89	20.35	15.04	19.47	11.50
		50	38.94	21.24	31.86	15.93	29.20	15.93
		100	38.05	30.09	23.89	19.47	22.12	19.47
		200	43.36	29.20	28.32	18.58	28.32	18.58
		400	46.90	32.74	34.51	23.01	34.51	23.01
		600	44.25	33.63	33.63	22.12	33.63	22.12
		800	46.90	36.28	32.74	26.55	32.74	26.55
		1000	49.56	37.17	30.97	27.43	30.97	27.43
		1500	45.13	37.17	31.86	26.55	31.86	26.55
	10	10	25.66	24.78	7.96	14.16	7.96	9.73
		20	30.09	23.89	21.24	15.04	20.35	11.50
		50	41.59	22.12	34.51	16.81	31.86	16.81
		100	40.71	30.97	24.78	20.35	23.01	20.35
		200	45.13	29.20	30.09	18.58	30.09	18.58
		400	47.79	33.63	35.40	22.12	35.40	22.12
		600	43.36	36.28	34.51	24.78	34.51	24.78
		800	46.90	36.28	32.74	26.55	32.74	26.55
		1000	49.56	36.28	30.97	26.55	30.97	26.55
		1500	46.02	37.17	32.74	26.55	32.74	26.55
	25	10	28.32	24.78	9.73	14.16	9.73	9.73
		20	33.63	23.01	23.01	15.93	22.12	12.39
		50	38.94	23.01	33.63	17.70	30.97	17.70
		100	42.48	30.09	28.32	21.24	26.55	21.24
		200	44.25	30.09	32.74	19.47	32.74	19.47
		400	48.67	32.74	43.36	23.01	43.36	23.01
		600	46.02	34.51	40.71	24.78	40.71	24.78
		800	50.44	36.28	36.28	26.55	36.28	26.55
		1000	52.21	36.28	33.63	26.55	33.63	26.55
		1500	49.56	37.17	38.05	28.32	38.05	28.32
	50	10	30.09	23.89	12.39	13.27	13.27	8.85
		20	35.40	23.01	26.55	17.70	25.66	14.16
		50	41.59	23.89	34.51	20.35	31.86	20.35
		100	41.59	29.20	30.97	23.89	29.20	23.89
		200	43.36	29.20	37.17	27.43	37.17	27.43

		400	51.33	29.20	46.02	23.01	46.02	23.01
		600	46.90	30.97	43.36	21.24	43.36	21.24
		800	46.02	38.05	42.48	28.32	42.48	28.32
		1000	46.90	37.17	42.48	25.66	42.48	25.66
		1500	46.02	38.05	43.36	29.20	43.36	29.20
	100	10	30.09	23.01	11.50	12.39	11.50	7.96
		20	38.94	23.89	30.09	18.58	29.20	15.04
		50	46.90	23.89	41.59	22.12	38.94	22.12
		100	40.71	30.97	38.05	27.43	36.28	27.43
		200	42.48	31.86	39.82	31.86	39.82	31.86
		400	48.67	30.97	48.67	26.55	48.67	26.55
		600	47.79	31.86	46.02	25.66	46.02	25.66
		800	45.13	37.17	46.02	29.20	46.02	29.20
		1000	43.36	36.28	45.13	30.09	45.13	30.09
		1500	44.25	36.28	42.48	30.97	42.48	30.97
	200	10	35.40	32.74	22.12	21.24	22.12	16.81
		20	40.71	27.43	35.40	22.12	34.51	18.58
		50	45.13	23.89	41.59	22.12	38.94	22.12
		100	40.71	32.74	41.59	31.86	39.82	31.86
		200	41.59	35.40	38.94	33.63	38.94	33.63
		400	49.56	36.28	51.33	33.63	51.33	33.63
		600	44.25	35.40	47.79	30.97	47.79	30.97
		800	46.02	38.05	49.56	31.86	49.56	31.86
		1000	44.25	38.05	46.90	31.86	46.90	31.86
		1500	46.02	37.17	46.90	31.86	46.90	31.86
	500	10	44.25	32.74	34.51	20.35	34.51	15.93
		20	43.36	30.09	37.17	23.01	36.28	19.47
		50	50.44	24.78	43.36	21.24	40.71	21.24
		100	40.71	31.86	48.67	33.63	46.90	33.63
		200	46.02	32.74	43.36	30.97	43.36	30.97
		400	47.79	34.51	49.56	33.63	49.56	33.63
		600	44.25	34.51	51.33	31.86	51.33	31.86
		800	46.90	37.17	52.21	30.97	52.21	30.97
		1000	51.33	40.71	52.21	34.51	52.21	34.51
		1500	48.67	38.94	53.10	33.63	53.10	33.63
	1000	10	44.25	33.63	32.74	23.01	32.74	18.58
		20	40.71	30.97	34.51	25.66	33.63	22.12
		50	46.90	26.55	41.59	21.24	38.94	21.24
		100	47.79	31.86	51.33	30.97	49.56	30.97

	200	44.25	34.51	36.28	30.97	36.28	30.97
	400	46.90	36.28	48.67	33.63	48.67	33.63
	600	45.13	36.28	52.21	33.63	52.21	33.63
	800	48.67	38.94	53.98	34.51	53.98	34.51
	1000	48.67	41.59	53.10	37.17	53.10	37.17
	1500	46.02	39.82	55.75	36.28	55.75	36.28