# COMPARATIVE ANALYSIS OF ENTROPY ALGORITHMS TO DETERMINE THE MOST EFFECTIVE TECHNIQUE FOR MEASURING COMPLEXITY IN BUILDING CONSTRUCTION

A Dissertation

by

CHAO XIAO

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

DOCTOR OF PHILOSOPHY

| | |
|---|---|
| Chair of Committee, | José L. Fernández-Solís |
| Co-Chair of Committee, | Rodney C. Hill |
| Committee Members, | Zofia. F. Rybkowski |
| | John A. Walewski |
| Head of Department, | Ward V. Wells |

August 2015

Major Subject: Architecture

ABSTRACT

Scholars have indicated that construction operation inefficiency is due to particular complexity factors owing to industry specific uncertainties and interdependences. The study of complexity in construction has become an essential topic to provide advanced methods and concepts for construction industry. It also has raised valid questions: Is construction really complex or just complicated? More importantly, how to measure the complexity in building construction systems?

This dissertation is based upon these two questions, and intend to fill the research gap that no quantitative complexity measurement has ever been found in research works. Comprehensive literature search is firstly used to make an embedded conceptual analysis of basic concepts of complex and complicated, to conclude building construction systems as complex systems and to metonymic map complex to construction domain. Chaos theory was then used to linked complex building construction systems and entropy complexity measurement together and proposed to use entropy algorithms to measure complexity in building construction.

However, entropy in construction could be measured in multiple ways with different results. Therefore, three commonly used entropy algorithms, which are Approximate Entropy, Sample Entropy and Permutation Entropy, were compared along with Six Sigma Analysis and Maximal Lyapunov Exponent based on ten (10) pilots cases and

their simulated cases. Two Rounds of simulation were conducted using Monte Carlo Simulation by MATLAB in order to generate more random number to represent different circumstances in building construction performance associate with different sample sizes.

The outcomes indicated that the compared with Approximate Entropy and Permutation Entropy, the characteristics of Sample Entropy make be sensitively and efficiently to tell different construction performance circumstances apart by significant complexity measurement for either small sample or large sample. This quantitative measurement of complexity in building construction not only fill the knowledge gap; it also avoids the subjectivity of evaluators and set a unified standard for complexity measurement in building construction in the future research.

Understanding complexity in construction management is important for two reasons: (1) to visualize how both complicated and complex traits exist in a construction project (object and social systems), and (2) to identify for stakeholders new types of managerial competencies and tools that reflect the understanding of complexity in construction.

## DEDICATION

*To my father Zhanjun Xiao and my mother Qinghong Yan*

## ACKNOWLEDGEMENTS

I would like to thank you my committee Chair, Dr. José L. Fernández-Solís, my co-Chair, Prof. Rodney C. Hill, and my committee members, Dr. Zofia. F. Rybkowski and Dr. John A. Walewski, for their guidance and support through the course of this research.

Thanks also go to my colleagues, my friends, especially my boyfriend Yang Liu and my friends Dr. John M. Nichols, Dr. Darrell E. Thompson and Mr. Mike Duff, and the department faculties and staffs for making my time at Texas A&M University a great experience.

Finally, thanks to my parents for their encouragement, support and love.

## NOMENCLATURE

| | |
|---|---|
| ApEn | Approximate Entropy |
| PEn | Permutation Entropy |
| PPC | Percent Plan Complete |
| SampEn | Sample Entropy |
| Std | Standard Deviation |
| TFV | Transformation-Flow-Value Theory |

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# CHAPTER I

# INTRODUCTION

## *1.1 Research Background*

Current research indicates that construction is approximately 40-60% as efficient as the manufacturing industry (Dubois & Gadde, 2002). Solving this low efficiency and high wastage problem is a priority for construction managers and researchers. Among all the studies, one conclusion is that construction operation inefficiency is due to particular complexity factors owing to industry specific uncertainties and interdependences. Winch (1989) indicated that construction projects are amongst the most complex of all undertakings. Gidado (1996) further emphasized this view by stating that there is a continuous increase in the complexity of construction projects. In construction, there are a number of sources for the resources employed, the environment in which construction takes place, the level of scientific knowledge required, and the number and interaction of different parts of the workflow. Dealing with complexity is one of the most difficult, but also the most critical, issues to improve construction performance (Baccarini, 1996; Gidado, 1996; Bertelsen, 2003).

The construction industry has its own system properties, which are much different from manufacturing as summarized in Table 1 In the construction industry, the design and production of unique and complex projects in highly uncertain environments under great time and schedule pressure is fundamentally different from making industrial products (Gregory, 1999). In addition, it has its own type of organization. The most common method for forming a temporary organization in the construction industry is general contractor plus sub-contractors. According to Fernandez-Solis' (2013) pioneering study, the physical production procedure in a construction project is the one that produces the final building output. Only subcontractors and their labors are directly involved in this process; they have their own project number and achieve the conversion from budgets, materials and time to building products. However, other stakeholders including owner, designer, even general contractor, who do not engage in this physical production procedure, cannot direct control the work of the field (Rounds & Segner, 2010). Rather, they deal with the information of sub contractors' projects. This project of projects is named a meta-project by Fernandez-Solis (2013).

Table 1 Comparison between Manufacturing Products and Construct Products

| **ITEMS** | **MANUFACTURING PRODUCTS** | **CONSTRUCTION PRODUCTS** (Fernandez-Solis, 2008) |
|---|---|---|
| Operation Organization | ▪ A fixed production enterprise, which is a long-term stable organization. | ▪ A temporary organization for a particular project within a limited duration. |
| Materials | ▪ Repeated homogeneous productions.<br>▪ Materials and components are quantitative and modeled to improve the efficiency. | ▪ Each project is a one-time production with different materials, components, structures and systems. |

Table 1 Continued

| ITEMS | MANUFACTURING PRODUCTS | CONSTRUCTION PRODUCTS (Fernandez-Solis, 2008) |
|---|---|---|
| Site | ▪ Made in stationary production workshop, thereby reducing the environmental uncertainty. | ▪ Each project is on a unique site with its own conditions, constraints and benefits. |
| Duration | ▪ Similar products have fixed production duration, which mainly depends on the produce technology | ▪ A different time limit due to constantly changing economic and natural conditions and circumstances. |

In a meta-project system, the general contractor and sub-contractors make their own promises about what they will achieve in a certain time period. General contractors actually manage promises. Simply keeping promises and conducting each task on time makes a successful project. However, due to the unique complex nature of the construction organization and its performance, breaking promises frequently occurs. This not only causes waste for certain participants; it also affects other participants' performance (Fernandez-Solis, 2013). Furthermore, the complex nature of construction also causes another kind of waste called systemic waste. Due to all of these, the complex nature of building constructions becomes the main reason for the high wastage and also for inefficient construction performance. This also becomes the motivation of complexity research in building construction.

### *1.2 Research Hypothesis and Research Question*

The first task for complexity research is to define the research objectives. There is often conflict over the nature of building construction systems, whether they are characterized as mainly complex or just complicated. To explore this, the first hypothesis states that building construction systems should be considered complex systems. Based on this assumption and its verification, most existing research about complexity in the construction field has focused on finding evidence for construction complexity and understanding it from different perspectives (Baccarini, 1996; Gidado, 1996; Bertelsen, 2003). Because of this, for most construction managers, complexity is still only a rough idea and is seldom used in real work (Curlee & Gordon, 2011). In order to solve this problem, a comprehensive analysis of complexity in building construction is necessary, especially if we could find a quantitative method to present complexity more objectively and intuitively. For this purpose, the following hypothesis starts from chaos theory to introduce an entropy model as a quantitative measurement of construction complexity. It tests whether it would help construction managers understand complexity and promote further application of complexity theory in building construction so as to improve construction performance.

Entropy describes the volume of change. It is one of the most important concepts in chaos theory and also in the proposed research. Since it was developed in the 1850s, it has expanded its definitions and application in three major fields: describing "waste heat" in physics and chemistry; measuring "disorder" in statistical views; and describing an analogous loss of data in information transmission systems (information entropy). In building construction, the most commonly used definition of entropy is the third one, which is to quantify information uncertainty in science domains or to evaluate organizational structure from the perspective of information transformation. No one has ever tried to consider entropy as an effective measure of construction complexity, even though it has been explored in several entropy models and has been widely used in other fields, such as information technology, physiology, business and health care. As a result of this, the main objective of the proposed research is a scientific search for new knowledge: could an entropy model be the most effective technique for measuring complexity in construction?

### *1.3 Research Outline*

In order to fill the research gap, and link the entropy model and building construction, the following structure and methodologies for the research work will be: 1) use a literature review for a full analysis of the complexity in building construction systems, which includes a comprehensive and embedded theoretical model of complexity science; 2) based on pilot studies and simulated comparative analysis among different kinds of entropy previously applied in complexity measurement, find out the most understandable and practical mathematical model to calculate the complexity of building construction systems; and showcase how it could benefit construction managers working on real projects.

# CHAPTER II

# LITERATURE REVIEW

## *2.1 Literature Search Approach*

The topic of entropy came from the research topic of Dixit (2010), which explored an embodied energy measurement for a building project. In the field of energy research-thermodynamics, entropy is associated with energy. Entropy is the degree to which energy is wasted (dispersed) and also a measure of the unavailability of heat energy for work, or so-called disorder (Clausius, 1867). Therefore, the first exploration for the dissertation is whether or not embodied entropy is present in a building construction project. “Entropy” appears as a keyword in the literature searches of several popular databases in civil engineering and the construction field. In order to connect entropy models with complex nature of building construction by applying chaos theory, keywords like “complex/complexity”, “chaos/chaotic” and “building construction/construction project” are also combined with “entropy” together for literature exploration. The main databases used in this literature search were ASCE Library, Compendex (Ei Village 2), Web of Knowledge, GreenFILE (Ebsco), ScienceDirect, ProQuest Dissertations and Google Scholar.

## *2.2 Summary of Literature Search*

### 2.2.1 Types of entropy

Based on entropy research in various fields, the topics could be divided into three major categories, according to Koskela's Production Theory (2000): transformation, flow and value. In the transformation category, entropy is defined in the usual way as the disposal of physical energy. In the flow category, entropy serves two different purposes depending on the field: some studies of chaos theory use entropy as a statistical concept to measure disorder or complexity, while information entropy is defined as the number of identical microstates. Finally, for the value view, entropy is often applied to assist in decision-making or measure the state of enterprises or organizations.

Entropy, as the measurement for energy disposal in transformation, is a fundamental application. However, the current engineering and construction research trend is to extend modeling entropy to the flow and value categories. Koskela's (2000) framework for construction incorporates all three categories into a systematic management. As a result of this, the literature is more numerous in the direction of information entropy and entropy in chaos theory. The literature matrix in Table 2 illustrates the studies that were explored, based on their references and citations.

Table 2 Literature Matrix Example

| Title | What is entropy | Related to construction? | Ref. 1 | Ref. 2 | Ref. 3 |
|---|---|---|---|---|---|
| Gößling-Reisemann, S. (2008). What is resource consumption and how can it be measured?. Journal of Industrial Ecology, 12(1), 10-25. (Has been cited 15 times) | Y | Y | P6.1: Gößling-Reisemann, S. (2008). Entropy analysis of metal production and recycling. Management of Environmental Quality: An International Journal, 19(4), 487-492. | P6.2: Ingwersen, W. W. (2011). Emergy as a life cycle impact assessment indicator. Journal of Industrial Ecology, 15(4), 550-567. | P6.3: Sulkowski, A., & White, D. (2009). Consumption of Energy, CO2 Emissions and Materials Usage Efficiency Per Capita: A Cluster Analysis of Europe and Eurasia. Global Management Journal, 1(1), 55-65. |
| Rödder, W. (2000). Conditional logic and the principle of entropy. Artificial Intelligence, 117(1), 83-106. (Has been cited 47 times) | Y | N | S1.1: Rödder, W., Gartner, I. R., & Rudolph, S. (2010). An entropy-driven expert system shell applied to portfolio selection. Expert Systems with Applications, 37(12), 7509-7520. | S1.2: Pankratz, G. (2005). A grouping genetic algorithm for the pickup and delivery problem with time windows. OR spectrum, 27(1), 21-41. | S1.3: Calabrese, P. G. (2005). Reflections on logic and probability in the context of conditionals (pp. 12-37). Springer Berlin Heidelberg. |
| Choi, J., & Russell, J. S. (2005). Long-term entropy and profitability change of United States public construction firms. Journal of Management in Engineering, 21(1), 17-26. (Has been cited 22 times) | Y | Y | D6.1: Christodoulou, S. E. (2008). A bid-unbalancing method for lowering a contractor's financial risk. Construction Management and Economics, 26(12), 1291-1302. | D6.2: Christodoulou, S., Ellinas, G., & Aslani, P. (2009). Entropy-based scheduling of resource-constrained construction projects. Automation in Construction, 18(7), 919-928. | D6.3: Christodoulou, S. E., Ellinas, G. N., & Aslani, P. (2009). Disorder considerations in resource-constrained scheduling. Construction Management and Economics, 27(3), 229-240. |

### 2.2.2 Information within Building Construction

Not only information entropy is associated with the average unpredictability in a random variable. This information perspective also could be found in the understanding of building construction systems. That is, the construction managers, especially general contractors, are actually managing information about all aspects of construction rather than being in control of the work in the field (Rounds & Segner, 2010).

The owner, designer, the general contractor, subcontractors and its team of consultants are all involved in the task of creating and managing information (what, when, where, how and why) that can be executed in the field. The range of study is marked in Figure 1.

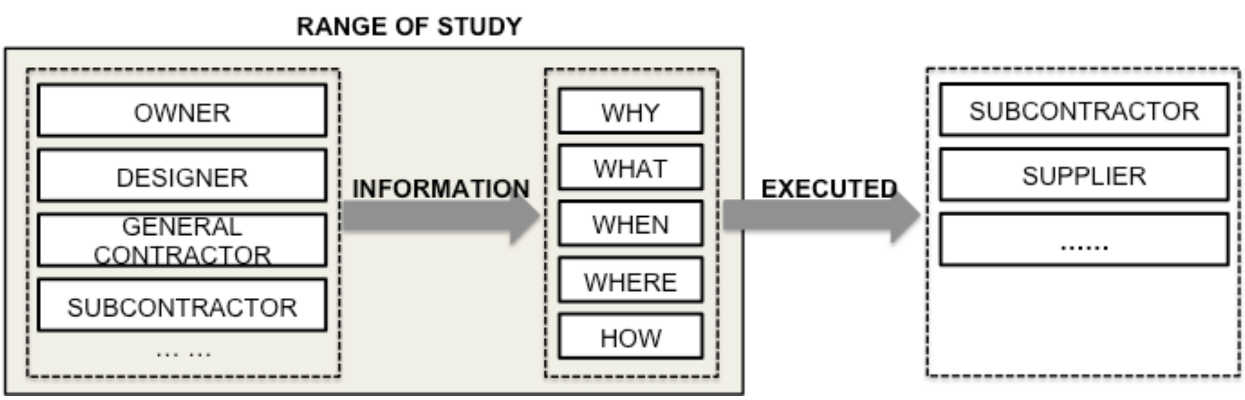



Figure 1 Construction System and Study Range

### 2.2.3 Complexity in Building Construction

Among all the literature, the paper of Baccarini (1996) was the most cited (cited by 413) and caught our attention. According to Baccarini, in construction, different sources and types of information, such as outside information, inside information, false information and asymmetric information, make the construction system more and more disordered and complex. Inspired by this idea, the second round of the literature search focused on construction complexity by using "complicated/complex/complexity," "chaos/ chaotic," "construction" and their combination as keywords. The studies indicate evidence for chaos and complexity in building construction, the preview and current research of construction complexity, and most importantly, the missing quantitative complexity-measuring model for construction practice.

In order to find a method for complexity calculation, the literature search also returned to entropy in disorder measuring, but this time, combining complex measure and entropy together as keywords. Three different kinds of commonly used entropy were found as complexity measurements in chaos theory. They are approximate entropy, sample entropy and permutation entropy. However, none of these entropy models has ever been used in construction systems for complexity measuring, and this is the research gap found based on the literature review.

This gap is investigated using chaos theory, which links an entropy model with the construction system to find out the most appropriate technique to measure construction complexity. The literature review map and the approach to finding the final proposal research topic is presented in Figure 2. The following parts summarize the literature reviews from several main aspects according to the keywords.

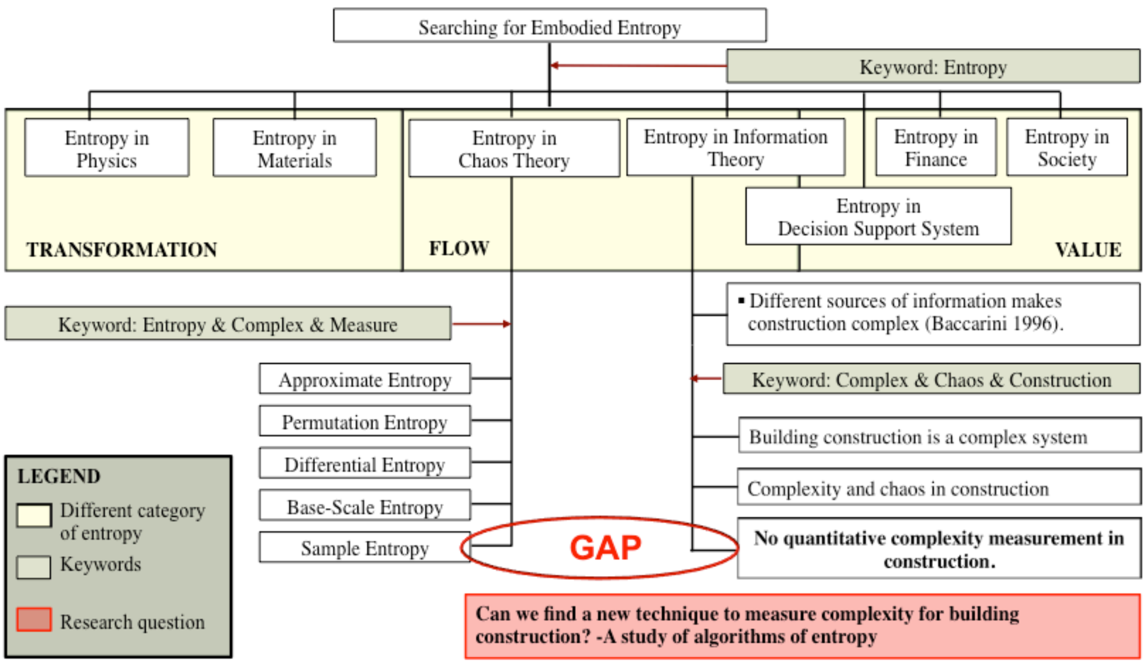



Figure 2 Literature Review Path

*2.3 Entropy*

Entropy, a word originating from Greek, describes the volume of change. It is one of the most important concepts in the chaos theory and also in the proposed research. It is the approach used in this study to measure complexity in building construction. As Wiener (1948), the originator of control theory, pointed out: the entropy of a system is the measurement of complexity and chaos.

### 2.3.1 Definition and Application of Entropy

The concept of entropy was used as early as the 1850s. With nearly two centuries of development, its definition and applications have expanded to more and more subjects. Among them, three basic understandings of entropy have been cited most. Originally, entropy described the "waste heat," or more specific, losing of energy, from heat engines and other mechanical devices, which, according to the Second Law of Thermodynamics, could never be 100% efficient (Rudolf, 1850). Later, the term's descriptions gradually expanded, as more was understood about the behavior of molecules on the microscopic level. Boltzmann, in 1872, proposed a microcosmic definition of entropy during his research on the movement of gas molecules. The word "disorder" was used in the study of statistical perspective of entropy, by applying probability theory into the description of the increased molecular movement on the microscopic level descriptions; entropy on the level of microscopic is resulted from statistical thermodynamics and its mechanics and then expanded in chaos theory. Since the mid-20th century, the concept of entropy has found applications in the field of information theory. Shannon (1948) developed the very general concept of information entropy to describe an analogous loss of data in information transmission systems. All the other ideas of entropy developed according to these three fundamental definitions, which could be summarized as classical thermodynamic view, statistical view and information theoretical view.

Entropy has been applied to more and more subjects; those related to construction and building production, along with some key studies, are categorized into three different perspectives according to the Transformation-Flow-Value (TFV) theory, as shown in Figure 3 Even though current practice in construction is based on the transformation view, the results of some prior research on contemporary construction and lean construction show that the future trend of construction management lies in the flow and value views (Koskela 1992, 1999, 2000). As a result of this, the second round of literature search mainly concentrated on two perspectives, entropy in chaos theory and entropy in information theory.

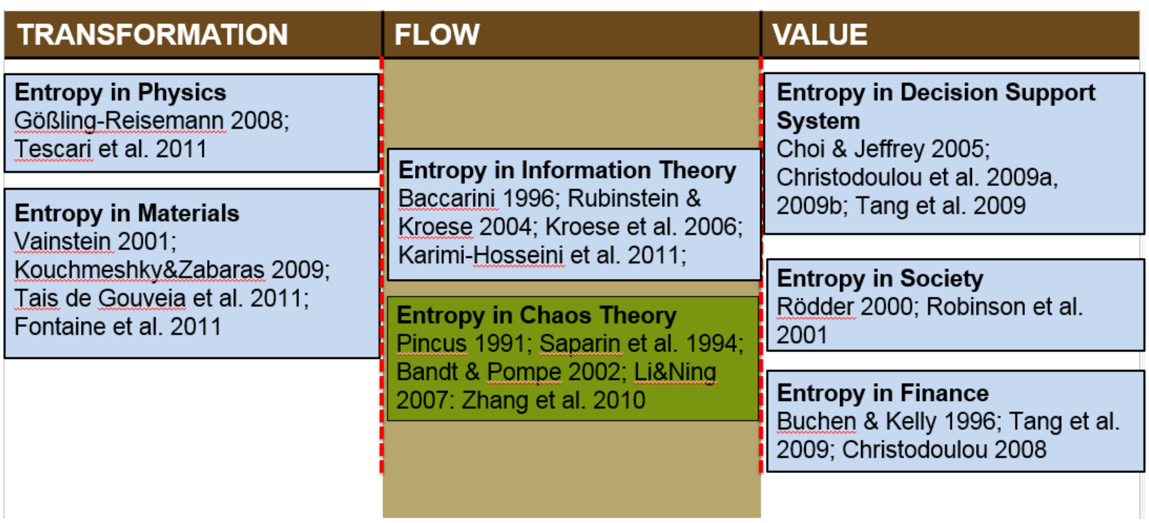



Figure 3 Summary of Related Entropy Research According to TFV Theory Based on Koskela (2000)

### 2.3.2 Entropy in chaos theory

In Boltzmann's definition (1872), entropy indicates how chaotic the particles' positions are in the system. As thermal energy always flows spontaneously from regions of higher temperature to regions of lower temperature, these processes reduce the state of order of the initial systems, and therefore entropy is an expression of disorder or randomness (Myat et al., 2012). The more chaotic the system, the greater its entropy and vice versa (Mao et al., 2009).

Certain papers concluded the existence of deterministic chaos from data analysis (Babloyantz & Destexhe, 1988) and included "error estimates" on dimension, the Lyapunov spectrum and entropy calculations (Zbilut et al., 1988). The classification of dynamical systems via entropy and the Lyapunov spectrum stemmed from the work of Kolmogorov (1963), Sinai (1959), and Oseledets (1968), though their works rely on ergodic theorems, and the results are applicable to probabilistic settings. Dimension formulas are motivated by a construction in the entropy calculation and generally resemble Hausdorff dimension calculations, which is a measure of the local size of a set of numbers by taking the distance between each of its members into account.

The entropy concept, as a measure of the degree of disorder in a system, is an indicator of a project's tendency to progress out of order and into a chaotic condition and it can thus serve to forecast a project's performance and its further development (Christodoulou et al., 2009a). Pincus (1991) and others' valuable research works apply different ideas of entropy as measure of system chaos and complexity in physics and systems engineering (Pincus, 1991; Saparin et al., 1994; Bandt & Pompe, 2002; Li et al., 2007: Zhang et al. 2010). The summary of each definition is listed in Table 3 for further comparison.

These three definitions of entropy and their correlated research papers have been cited in plenty of fields for complexity measuring. Approximate entropy and sample entropy are mainly applied in the fields of information technology, physiology and health care (Kurths et al., 1995; Richman & Moorman, 2000; Zhang et al., 2008; Khandoker et al., 2008). Permutation entropy is used in statistical and physical fields to distinguish noise from experimental data (Daw et al., 2003; Cao et al., 2004; Xu et al., 2008). However, none of the presented entropy estimators or other similar entropy concepts have been used in the building construction field to measure complexity.

Table 3 Summary of Entropy Models for Complexity or Chaos

| Entropy Model | Definition | Author & Year |
|---|---|---|
| Approximate Entropy (ApEn) | ApEn is a technique used to quantify the amount of regularity and the unpredictability of fluctuations over time-series data. | Pincus (1991) (Has been cited by 2859 times) |
| Sample Entropy (SampEn) | It is a modification of ApEn. It used extensively for assessing the complexity of a physiological time-series signal. But it does not include self-similar patterns as ApEn does. | Grassberger (1988); Richman and Moorman (2000). (Has been cited by 2184 times) |
| Permutation Entropy | It is an appropriate complexity in the presence of dynamical and observational noise. | Bandt and Pompe (2002) (Has been cited by 838 times) |

### 2.3.3 Entropy in Information Theory

Information entropy is well defined in the literature. Generally, it quantifies the expected value of the information contained in a message. This concept of information entropy was introduced by Shannon (1948) –"A mathematical theory of communication." This is the most influential paper about information entropy, having been cited 63187 times.

One important purpose of entropy used in construction is to quantify information uncertainty in science domains and to solve problems associated with information transmission. For example, in Eshragh's (2011) work, he proposed an adaptation of the Cross Entropy method called Projection-Adapted Cross Entropy to solve a transmission expansion problem that arises in management of national and provincial electricity grids.

The Cross Entropy method for solving rare event probability estimation was pioneered in 1997 (Rubinstein, 1997), and later expanded to solve combinatorial optimization problems (Rubinstein, 2004). Karimi-Hosseini also chose “transinformation entropy” as her main evaluation criteria for site selection of rain-gauges in the Gav-Khuni basin rainfall network (Karimi-Hosseini, 2011). In her research work, information transfer (transinformation) of random and continuous variables Z between two locations i and j in a monitoring is called transinformation entropy.

In the construction management filed, entropy is not only a useful tool to evaluate information transmission. Based on the theory of information entropy, an information entropy model of organizational structure is also constructed according to the characteristics of information flow in the execution of the construction program (Mao et al., 2009). If there is direct connection between elements X and Y of a construction organization, information can be sent from one to the other directly, and vice versa. Due to the uncertainties involved in project management information, any delivered information could be valid, true or false with a probability of $p_n$. Thus, the quantified uncertainty in the information communication from X to Y is information entropy, which could help determine the organization’s optimization. The entropy effect of project communication is mainly manifested in the transfer of communication effectiveness between each level and member. Even though excellent communication is crucial for a good construction organization and the success of a project (Zheng & Li, 2009), only a few successful cases were found in the literature research. As Baccarini

(1996) stated, the reason for this failure is mainly because of the complexity in construction projects; and most important factor in the success of a project is the management of complexity.

Baccarini's research (1996) of project complexity has been cited 417 times and inspired the following research of complexity in construction. It also guided our second round of literature search, and raised a question of whether the idea of entropy could be related to the management of complexity.

### *2.4 Complexity/Complex Theory*

Two fundamental studies have led the study of complexity in building construction: "The Concept of Project Complexity – A Review" (Baccarini, 1996), published in the *International Journal of Project Management*, and "Complexity – Construction in A New Perspective" (Bertelsen, 2003) in the *11th Annual Conference in the International Group for Lean Construction*. Both of these papers are cited in the research of complexity in construction (Williams, 1999; Del Cano et al., 2002; Abdelhamid, 2004; Vidal & Marle, 2008; González et al., 2008); at the same time, they refer to those basic research works in the concept of complexity (Bennett & Fine, 1980; Gidado, 1996; Radosavljevic et al., 2002). As a result of this, these two papers helped us to expand the literature review for complex studies not only backward but forward as well. The following sections summarize the results of complex studies and their applications in building construction.

“System” commonly means a group of interacting, interrelated, or interdependent elements forming a holistic functional whole (Crawford et al., 2005). However, the “complex” nature of systems elicits multiple definitions. This proposal summarizes the literature of complex system from two perspectives: their different levels of definition, and their dimensions.

No consensus exists with respect to a definition for “complex system,” as well as for the definition of complexity (Vidal et al., 2011). Complexity has different connotations within the same field. Different levels of complexity are expressed in Figure 4. The first definition is that complex systems integrate multiple thematic domains. None of the agents or subsystems could be fully understood when considered in isolation (Cilliers, 1998). The second and more detailed level sees complexity as emerging from nonlinearities due to the large number of interactions involving feedback occurring at one or more lower levels within the system (Batten, 2001; Manson, 2001). One more idea of “complex system” extends the notion of complexity by creating more refined representations of micro-level heterogeneity and interactive processes and factoring in top-down (perhaps emergent) structures that feed back to influence bottom-up phenomena (Parker et al., 2003).

An exploration of the literature also reveals a wide range of factors that may contribute to project complexity. These contributing factors are defined by Remington et al. (2009) in terms of dimensions; see Table 4.

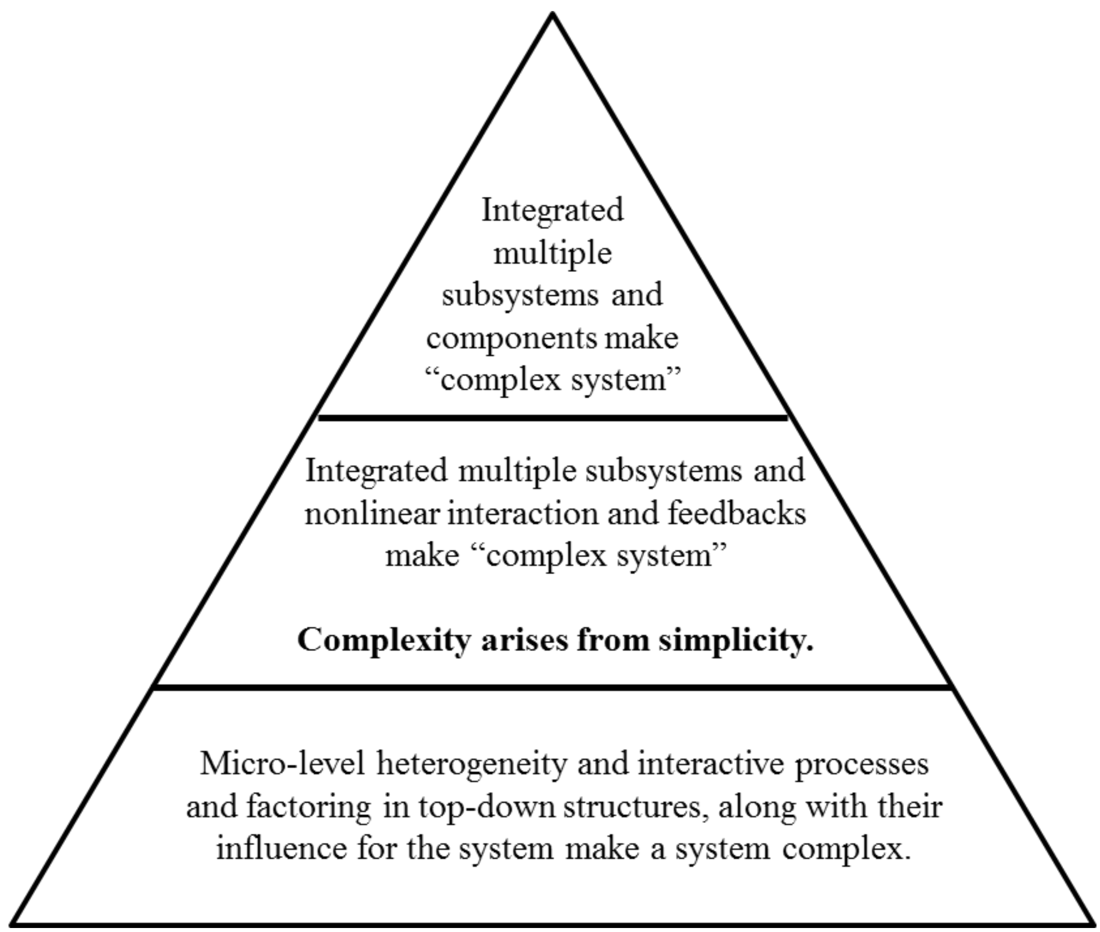



Figure 4 Three Different Understandings of Complex Systems

Table 4 Dimensions of Project Complexity (Adapted from Yugue & Maximiano, 2012)

| DIMENSION OF COMPLEXITY | AUTHORS |
|---|---|
| Uncertainty about the product | Turner & Cochrane (1993); Williams (1999); Remington, et al. (2009) |
| Uncertainty about the scope of the project | Turner & Cochrane (1993); Baccarini (1996); Tatikonda & Rosenthal (2000); |
| Novelty of technology | Baccarini (1996); Williams (1999); Geraldi (2007); Fitsilis (2009); Remington, et al. (2009); Tatikonda & Rosenthal (2000) |
| Highly multidisciplinary | Baccarini (1996); Geraldi (2007); Geraldi & Adlbrecht (2007); Fitsilis (2009) |
| Large number of stakeholders with influence on the project | Williams (1999); Fitsilis (2009); Remington, et al. (2009) |
| High difficulty to achieve performance goals | Remington, et al. (2009) |
| Significant change in the scope of the project during its implementation | Turner & Cochrane (1993); Williams (1999); Geraldi (2007); Geraldi & Adlbrecht (2007); Fitsilis (2009); Remington, et al. (2009); |
| High interdependence between the technologies | Baccarini (1996); Williams (1999); Geraldi (2007); Fitsilis (2009); Remington, et al. (2009); Tatikonda & Rosenthal (2000) |
| High interdependence between firms involved in the project | Baccarini (1996); Williams (1999); Geraldi (2007); Remington, et al. (2009) |

### 2.4.1 Is the building construction system characterized as mainly complicated or complex?

Based on the definitions of projects and systems, projects including construction projects may be considered as systems (Boulding, 1956; Vidal & Marle, 2008; Vidal et al., 2011). However, there are two different types of systems: complicated and complex. This classification frames the first hypothesis of this proposed research: is construction characterized as complicated or complex?

Research has concentrated on the analysis and comparison of these two systems. Hill (1982) started the study of complicated system; the research work of Cilliers (1998, 2002, 2005), Amaral and Ottino (2004) and Heylighen et al. (2007) focus on complex systems rather than complicated ones. Meanwhile, Grabowski and Strzalka (2008), Page (2008), and Dekker et al. (2013) have devoted their research to comparing these different concepts, such as simple, complicated, complexity and uncertainty. However, there are no common definitions for complicated system and complex system. Several characteristics are agreed upon by all the above scholars; they are compared and summarized in Table 5.

Table 5 Comparisons Between a Complicated System and Complex System

| **COMPLICATED SYSTEM** | **COMPLEX SYSTEM** |
|---|---|
| Consisting of a large number of interacting components | |
| ▪ Specialized structures, deterministic switching | ▪ Structures for general use, non-deterministic switching |
| ▪ Algorithmic processing | ▪ Interactive processing |
| ▪ Fully understandable | ▪ No rules or formulas can capture the whole system |
| ▪ Static planning of performance, Mean Value Analysis | ▪ Dynamic planning of performance at the edge of chaos |
| ▪ Bounded resources | ▪ An open system with unbounded resources |
| ▪ Lack of memory (independence of processes) | ▪ Existence of memory (dependence of processes) |
| ▪ Simple feedback | ▪ Self-organization |
| ▪ Having best method to operate the system | ▪ No best plan due to a changing environment |

The building (final product, noun):

- Consists of a large (bounded) number of interacting components
- Is a specialized structure
- Has an operation that can be mapped and put in algorithms
- Is fully understandable
- Has static performance, and mean value analysis of its functions can be readily made
- Uses bounded resources, except for operation and maintenance; those can be considered un-bounded
- Provides simple feedback loops
- Contains a best method to operate the facility composed of all its systems

To build, as in the physical work of construction, which encompasses materials, equipment and labor, a verb that mobilizes labor and equipment to erect and assemble materials, is characterized by the following statements:

- Consists of an extremely large number of contributing components (mineral extraction, fabrication, transportation, general economy, assembly, erection, finishing, code compliance; the list is as big as the amount of granularity we seek to inform).
- Non-deterministic switching is due to the fact that the players are autonomous agents. Autonomous agents have particular strategic, logistic and tactical interests and therefore may withhold asymmetric information from other players.

This non-deterministic switching property of complex systems, is relative to deterministic switching. Deterministic switching could be evaluated using specific algorithms, which means it is foreseeable. However, in a construction system, the autonomy of each participant, along with its own strategies that induce a variety of actions, lead to behavior ranging from almost deterministic actions to chaos-like dynamics, which makes the switching non-determinable.

- Processing is interactive, due to the amount, quality and type of information that all the stakeholders must contribute, check, verify and approve to achieve the intended results.
- No rules or formulas can capture the whole production system because each production is unique, one of a kind, different and distinct in multiple aspects, starting with the fact that the team of stakeholders is temporary and it intervenes as needed.
- Dynamic planning of performance at the edge of chaos is due to the fact that the autonomous players' interventions are predicated by activities in other projects and are determined by the strategic plan of each stakeholder.
- If the process takes into consideration all the materials and players necessary to make the final product, a large portion of national or world economies would be included, as the process is energy and material intensive as well as labor intensive and information super intensive.
- Existence of memory is required so that each process does not require reinvention each time it is needed.

- Self-organization is essential for the production system, as it has no central control.
- Having no best plan, due to a changing environment, is apparent from the above mentioned characteristics of the construction process or project delivery system.

According to Fernández-Solís' (2008) summary of the systemic nature of construction, construction operation and organization is a deterministic dynamic, non-linear flow, in which an extremely large number of stakeholders are involved. The outputs of construction are not proportional to the inputs and the whole is different from the sum of its parts, where the sum of its parts is much larger than the final product. The building construction system must consider an open social system and also the inter-operability of each participant (working inside the company) as well as extra-operability of the participants (working with other companies). It is nested in a social system with a varying team where communications and cooperation are emergent phenomena in each project; this emergent nature also helps the system learn from itself and achieve its self-organized goals toward the completion of the design intent. All of these natures of building construction's project delivery system favor considering it as a complex system. In spite of the fact that there are numerous views on explaining the differences between complicated and complex systems, most scholars mentioned in this paper agree with the differences summarized in Table 6; the only characteristic common to complicated and complex systems is that both contain a great number of components. In this study, the definitions, descriptions and distinctions between complicated and complex are used to shed light on the concepts of building and building construction system. While a building (noun, i.e. final product) is a complicated system, constructing a building (verb, i.e. to produce and erect the design intent) is a complex social system, which emerges from a deep and extended network of interactions and interconnection. It requires further detailed analysis.

### 2.4.2 Building Construction as a Complex System

Scientists have attempted to understand construction systems using a reductionist approach in which the behavior of a system is represented as being an equilibrium mechanical interaction of its components. This equilibrium assumption views spatial distribution as optimal and stationary (Allen, 2008). That is to say, rigid traditional construction management focuses on order, structure and planning. However, the unknowns in construction systems are better handled by a flexible process that promotes openness, coincidence and serendipity. For this reason, the behavior of complex systems offers an appropriate set of concepts with which to begin a new reflection on human systems, especially construction systems (Allen, 2008). Unlike the mechanical systems of a bygone era, overall system behaviors are no longer exclusively deterministic. In this new point of view, in construction systems, non-equilibrium phenomena are much more critical and offer a new way of understanding structural emergence and organization in systems with many interacting individual elements. In this complex system, all powers are connected. This new attitude toward construction systems puts forth a number of characteristics to predict project outcomes, in order to control or manage the construction procedure (Remington et al., 2009). Homer-Dixon's work (2002) summarizes the characteristics of complex systems and we apply them to construction performance:

1) Multiplicity: the number of ways that could produce a certain state. Construction's stakeholders consist of owner, architects, engineers, and consultants, and especially, construction teams like general contractors, sub-contractors, vendors and suppliers. Each organization has an identical structure of strategic, logistic and tactical personnel.

2) Causal connections: numbers of links between components (to the extreme, there is causal feedback where a change in one component loops back to affect the originals). This is common in construction practices. One of the most common examples is the misunderstanding of a client's requirements by the design team, which causes project schedule delays or over-budget performance.

3) Interdependence: the larger the module that can be removed from the complex system without affecting the overall system's behavior, the more resilient and less complex the system. However, in the construction system, none of the tasks, parts or units involved in the process can be easily removed or replaced, which makes construction a highly complex industry (Gidado ,1996). Supply Chain and

contractors’ networks are good examples to illustrate interdependence in construction systems.

4) Openness: to outside environment, not self-contained, difficult to locate boundary. Construction systems are nested in a social environment. Policies, economic situations and advanced technologies regularly affect construction systems.

5) Emergence: the degree to which the entire system is more than the sum of its parts, because a system may transcend its components. This is also the philosophical core of complexity theory (Elnashaie & Grace, 2007). According to the recently specified “value” theory in construction (Koskela & Howell, 2002), a constructed project has no value until it is turned over to the owner and used for its intended purpose; the value of the whole building is greater than the sum of parts.

6) Nonlinear behavior: the effect on the system in not proportional to the size of the change on a component. The nonlinearity of the construction project carries through the nonlinearity of the subsectors, and the nonlinearity of the industry and the general economy. In construction projects, most mathematical equations are complex and nonlinear, and generating these equations can be a remarkable challenge (Azimi et al., 2012). For instance, the total cost is not only related to

the building mass; it is also affected by construction location, delivery method and so on.

7) Adaptiveness: Organizations are adaptive in that they do not simply respond to events or surrounding circumstances, but evolve or learn. Each component, or we can say agent, in the construction system is guided by its own rules of behavior and also by a schema shared with others in the system (Thomas & Mengel, 2008). Curlee and Gordon (2010) agree that transformation or adaptive-ness is the nature of construction projects. One of the most famous theories of this is the transformational leadership idea in lean construction.

### *2.5 Current complexity research in building construction*

This dynamic complexity of construction projects intuitively results from the stochastic spatial and temporal interactions among multiple components, such as on-site equipment resources, labor productivity, unexpected external events, and human decisions regarding resource allocation and activity rescheduling (Tang et al., 2010). However, this is not the real complexity for this long-term research. As Curlee and Gordon indicated in their book *Complexity Theory and Project Management* (2011) that complexity does not necessarily reflect complicated or large projects, nor does it imply technical difficulty. In fact, it is concerned with the behavior over time of certain complex systems, while the complex systems combine elements of both ordered and random behaviors in an elusive but striking manner (Stwart, 2002; Allen, 2008). From Thomas and Mengel's statement (2008), we can categorize three different levels of complexity research:  complex system, complexity theory and complexity science. The current research of complexity in building construction may also be summarized from these three levels.

### 2.5.1 Complex system

Current research treats complex construction/project systems as objects mainly. Results reveal that complex systems have numerous unique characteristics, which could be referred to to as construction performance as well (Gidado, 1996; Koskela, 2000; Koskela & Howell, 2002; Azimi et al., 2012). According to Homer-Dixon (2000), they can be summarized as multiplicity, causal connections, interdependence, openness, emergence, nonlinear behavior and adaptiveness. This new concept of a construction system makes it extremely difficult to predict project outcomes, to control or manage construction procedures (Remington et al., 2009).

### 2.5.2 Complexity theory

Complexity theory is a relatively new way of thinking about systems of interacting components, such as firms and projects. It rests on the idea that order emerges through the interactions of components or agents (Benbya & Mckelvey, 2006). Complexity theory suggests that attempts to rigidly control a complex system can increase problems and unintended consequences as individuals in the system “work around” these controls (Zimmerman et al., 1998). It also suggests that, in order to represent changes in a complex system, it is important to understand the recurring patterns in the system, including the patterns of interactive relationships (Cohen & Stewart, 1994). Research about complexity theory in construction (Turner & Cochrane, 1993; Baccarini, 1996; De Meyer et al., 2002; Williams, 2005) has proven that taking complexity theories seriously in the construction management field would (a) help us to understand the current construction project environment in new and different ways, and (b) require new types of competencies for contractors and their performance.

### 2.5.3 Complexity Science

Complexity science thinking within the natural sciences began in the 19th century with roots stretching back ,with at least of early work on cellular automata, cybernetics, and general systems theory（Spencer, 1887; Manson, 2001; Crawford et al., 2005). It continued into the 20th and 21st centuries, with scholars from politics, social policy, social network, geography and healthcare applying complexity science within their disciplines. As Linstone (1999) stated, complexity science was the most exciting development in the systems area in recent years. However, it is not easy to define “complexity science,” given its long gestation and continuing growth and maturation (O’Sullivan et al., 2006). However, a comparison between complexity science and established science, as shown in Table 6, is cited and accepted in the research about complex science.

Table 6 Complexity Science Compared with Established Science (Adapted from Begun et al., 2003)

| COMPLEXITY SCIENCE | ESTABLISHED SCIENCE |
|---|---|
| Holism | Reductionism |
| Indeterminism | Determinism |
| Relationship among entities | Discrete entities |
| Nonlinear relationship | Linear relationship |
| Critical mass thresholds | Marginal increase |
| Quantum Physics | Newtonian Physics |
| Influence through iterative nonlinear feedback | Influence as direct result of force from one object to another |
| Expect novel and probabilistic world | Expect predictable world |
| Understanding; sensitivity analysis | Prediction |
| Focus on variation | Focus on averages |
| Behavior emerges from bottom up | Behavior specified from top down |
| Metaphor of morphogenesis | Metaphor of assembly |

Manson (2001) reviewed a diverse literature of complexity and summarized three categories of research where the term is used: algorithmic complexity, deterministic complexity and aggregate complexity:

**Algorithmic Complexity**: refers to measurement of the difficulty of computational problems (O’Sullivan et al., 2006). The most widely known definition of algorithmic complexity is the Kolmogorov–Chaitin measure, which is often referred to simply as the Kolmogorov complexity. The Kolmogorov complexity of an object, such as a piece of text, is broadly defined as the length in bits of the shortest description for that object.

Alternatively, it is the length of the shortest program required to obtain the output (Dewey, 1997).

**Deterministic Complexity**: refers to the unpredictable dynamic behavior of relatively simple deterministic systems (Manson, 2001). According to this definition, unpredictability is framed as sensitive dependence of outcomes on initial conditions.

**Aggregate Complexity**: the study of phenomena characterized by interactions among many distinct components (Manson, 2001). This is the most comprehensive definition with encouraging characteristics for the analysis of building construction systems.

The notion of complexity has been widely studied in fields such as astronomy, chemistry, evolutionary biology, geology and meteorology (Nocolis & Prigogine, 1989). However, its translation into the project management field started in the 1990s. In the project management field, the science of complexity seeks systematic and deliberate reduction in order to harness chaos in a manner that allows the project manager to increase his/her team's effectiveness by allowing a certain degree of individuality to move a project forward (Allen, 2008).

### 2.5.4 Similarities and Differences between Complexity Theory and Science

In the last 30 years, in particular, there has been a re-evaluation of the nature of complexity, and more fundamentally, of the relationship between order and disorder (Serres et al., 1982). Both complexity theory and complexity science focus on the relationships between these elements rather than on each element alone within the system. In addition, they all target the complex system as mainly objective, and provide a perspective to organize and manage a system or a project not only by simply linear prediction.

Complexity theory is a new way of thinking about the complex systems; and dealing with its typically characteristics. It is concerned with the behavior over time of certain kinds of complex systems. Taking complexity theories seriously in construction management field would (a) help us to understand the current construction project environment in new and different ways, and (b) would require new types of competencies for contractors and their performance. Most authors like Turner and Cochrane (1993), Baccarini (1996), and Williams (2005) have tended to focus on uncertainty and difficulty of the technical or management challenges, or about the organizational complicacy.

Complexity sciences are a relatively eclectic collection of academic efforts crossing a wide variety of disciplines. It reframes our view of many systems, which are only partially understood by traditional scientific insights. Complexity science is not a single theory. It is the study of complex systems, including the patterns of relationships within them, how they are sustained, how they self-organize and how outcomes emerge. Within the science there are many theories and concepts. The science encompasses more than one theoretical framework, such as biologists, anthropologists, economists, sociologists, management theorists and many others in a quest to answer some fundamental questions about living, adaptable, changeable systems (Zimmerman et al., 2000). Complexity science describes how systems actually behave rather than how they should behave.

From the science perspective, complexity research from the project management discipline has produced a number of approaches:

- Turner and Cochrane (1993) first connected project complexity with lack of clarity on project goals.

- The first established dichotomy about complexity is from Baccarini (1996). He proposed the complexity of a project could be interpreted and measured in terms of differentiation and interdependencies for both organizational complexity and technological complexity. The differentiation holds two dimensions: vertical differentiation and horizontal differentiation.

- Based on previous studies, Remington and Pollack (2007) categorized complexity into four dimensions, which are factors that characterize the nature of the complexity or a mixture of the two, based on the source of complexity: structural, technical, directional and temporal.

- Vidal and Marle (2008) summarized the historical research of project complexity into two main scientific approaches. The first one, usually known as the field of descriptive complexity, considers complexity as an intrinsic property of a system. An example of this vision is the work of Baccarini (1996), which incited researchers to find methods and tools to quantify or measure complexity. The other one, usually known as the field of perceived complexity, treats complexity as subjective, seeing the complexity of a system as improperly understood through the perception of an observer. On the basis of this, a project manager deals with perceived complexity, as he/she cannot understand and deal with the full reality and complexity of the project.

All of the above research has emphasized that a clear understanding of complexity helps in selecting appropriate tools and approaches to manage a project successfully. However, all these existing works in the science of complexity, no matter what perspective it is, are theoretical research or qualitative research of complexity. The research gap we found is that quantitative research or measurement of complexity in building construction remains scarce. One possible support for this statement, according to Lu (2010), there is no fundamental conceptual framework with sufficient explanatory power has effectively integrated the mathematical model, complexity science and the practices of construction together until now. Furthermore, most scientists in the construction area still view complexity as an abstract concept, which is not familiar to them. As a result of this, complexity theory has not yet been considered as a practical method and an applicable tool, which is reflected in its limited applications in research studies of construction. In order to integrate the mathematical model and complexity science together and fill the research gap, chaos theory became a useful tool in the second hypothesis of the proposed research, which is using chaos theory as a bridge to link the complexity science in building construction systems and the entropy complexity measuring model together.

### *2.6 Chaos as a Representation of a more Complex Order*

The complexity of construction systems stems from potential non-linear, emergent behavior that can occur in interactions between interconnected tasks (Remington et al., 2009). Actually, chaos is where complexity arises through the non-linear interactions of small numbers of simple components and parts, people, equipment, materials and so on (Curlee & Gordon, 2011; Fernández-Solís et al., 2013, 2014). Project complexity is interested in the two zones to which a disturbed system may return: stable or unstable zones. Under appropriate conditions, systems may operate at the boundary between these zones, sometimes called a phase transition, or the “edge of chaos” (Rosenhead, 1998). Thompson and Gray (1990) provide the elements for a graphic metaphor on the range from order to disorder with its transitions. Based on their ideas, we summarized the different status of a system in a more vivid graphic metaphor, shown in Figure 5. In building construction systems, Fernández-Solís et al.'s (2013) case study of chaos arrived at a series of similar graphic images of PPC time series data, which verified our metaphor.

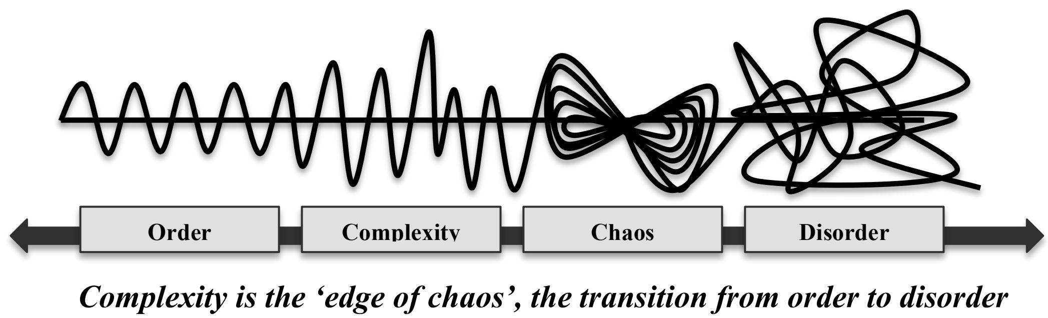



Figure 5 Graphic Metaphor of Order, Complexity, Chaos and Disorder

From Figure 5, we surmise that in a complex system, we can expect an orderly start that flows through and into transitions to complexity and then experiences chaotic or near chaotic episodes in an ambient prone to disorder. This conclusion fits well with the observation that the construction effort is a constantly focused effort at applying information and knowledge against an ever present tendency towards entropy. Entropy, the increase of disorder which is conquered through labor, nevertheless takes place in the arena, the ambience, of the universe of disorder. The complex construction system also has no long-term equilibrium, and just like other social systems, there is no historical evidence for long-run equilibrium; there is evidence only for the apparent "chaos" created by a complex interplay between a numbers of forces (Stevenson & Harmeling, 1990).

McKercher (1999) treats chaos and complexity as companions. His view is supported by a number of other researchers, such as Lewin (1993), Faulkner and Ressell (1997), and Byrne (1998). Chaos describes a situation where a system is dislodged from its balance and stability, which is random and unpredictable for outcomes. Complex systems often show nonlinear phenomena as chaos. Therefore, chaos theory is closely allied with complexity theory or the theory of complex science. Ineluctably, there has been debate about the differences between chaos and complexity. Axelrod and Cohen (1999) argued that chaos deal with situations, which lead to disorganized and unmanageable systems, while complexity theory deals with systems that have a large number of subsystems or elements and although hard to predict, these systems have structural and permit improvement. In spite of these disagreements and arguments, many see the close linkage between chaos and complexity theory (Lu et al., 2010). Waldrop first (1992) considered complexity as the emerging science at the edge of order and chaos.

Other similar expressions exist about their tight relationship. Curlee and Gordon (2011) concluded that complexity theory grew from chaos theory, and it was about harnessing chaos in a manner that allows the project manager to increase his/her team's effectiveness by allowing a certain degree of individuality to move a project forward. Chaos theory offers a solid theoretical and methodological foundation for interpreting a wide class of nonlinearity, instability and uncertainty, which characterize the increasingly complex systems. More detailed, chaos theory poses a promising and valid alternative theory for modeling complex, unsteady and unpredictable dynamics, which are associated with many aspects of construction and building systems. This makes chaos theory and chaos models attractive, since small models can offer real-world like phenomena which are fully determined by internal dynamics. It is worth mentioning that the system is dependent on interaction with its surroundings; this influence is precisely what nonlinear parameters try to describe. Nonlinear dynamic complex systems and chaos theory have great potential, not only from a research point of view, but also for suggesting new applications. When we combine complex building construction systems, chaos model based analysis and visualization, this may open a potential practical way to identify, forecast, control, and provide insight into complex behaviors in numerous kinds of systems, especially construction activity and organization systems.

The above explanations and arguments provide a possible way to fill the research gap, that is, we could use chaos as a modeling basis to validate and reflect the increasing complexity of complex construction systems, and to link the mathematical measurement model and complex construction performance together.

## *2.7 Conclusion*

The above literature exploration presents how the proposed research topic came about. In searching for "embodied entropy," we found the application of entropy as the measure of complexity, and the complex nature of building construction systems. Even though most of the studies point out the importance of considering building construction as a dynamic, non-linear, and emergent complex system, they carry out their research on only one part or perspective of complexity in building construction. There is no fundamental and widely accepted conceptual framework of complexity in building construction, which leads to a missing quantitative complexity measurement model for building construction. This has resulted in current ineffectual and unsuccessful construction management projects. Even though several different kinds of entropy have been used as effective complexity measurements, none of them have been introduced in building construction systems to analyze the complex nature quantitatively and to improve the management of construction.

# CHAPTER III

# METHODOLOGY

The research methodology encompasses each single method used that directly affects the subject and outcomes of this study. Integrating a comprehensive literature review, conceptual analysis, and qualitative and quantitative comparisons forms a mixed-method analysis. The reason for selecting mixed-method analysis for a single study is to gain a better understanding of the research problem and more persuasive results (Teddlie & Tashakkori, 2003).

## *3.1 Research Methods and Procedure*

In order to link the entropy model and building construction together, the study comprises three (3) phased procedures in the adapted mixed-method analysis. This multi-step mixed research procedure with both qualitative and quantitative methods is believed to provide a more comprehensive analysis than either approach could provide alone (Caracelli & Greene, 1997; Leedy & Ormrod, 2005). This multi-step research process offers a more rigorous analysis by incorporating the advantages of comprehensive literature review, conceptual analysis, and qualitative and quantitative comparisons.

The foundation for the study is a full analysis of the complexity in building construction systems, which includes a comprehensive and embedded theoretical model of complexity science. Among different kinds of entropy, which have already been applied in complexity measuring, an understandable and practical mathematical model to calculate complexity at each level would be based on in-depth analysis and the complex nature of building construction systems. To achieve the research objective, the research methodology was divided into three stages, graphically illustrated in the Figure 6, which are described briefly in the following:

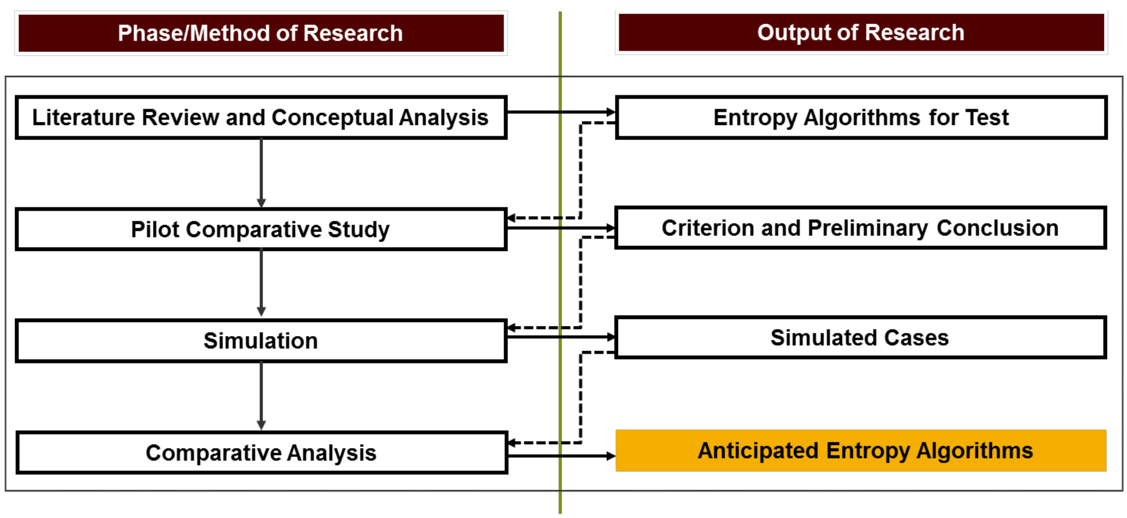



Figure 6 Research Procedures and Output

1. At the theoretical level, the research goal is to make an embedded conceptual analysis and summary of basic concepts in complexity science and chaos theory in order to pick up the most commonly used entropy models for complexity measurement for further comparison.

2. At the mathematical level, the research goal is to discover the best systematic and understandable mathematical model for complexity in construction based on a comparative analysis of algorithms of entropy. Different entropy methods applied in complexity measuring are compared on the theoretical level, and also through a pilot study of ten representative cases.

3. At the practical level, the research goal is to verify the results in real building construction projects, and to make sure the proposed quantitative measuring model could be accepted in real construction work. Monte Carlo simulation generates more numbers for each case in order to mimic real construction situations over a long period. Based on the simulated cases, the entropy algorithms are further compared, the complexity of construction is measured and also, different circumstances are tracked.

### *3.2 Literature Review and Conceptual Analysis*

In order to build the theoretical model of complexity in building construction as the foundation for the research and to select appropriate entropy algorithms for further analysis, a comprehensive literature review and conceptual analysis are the primary methods used in the research.

The databases used for the literature search include ASCE Library, Compendex (Ei Village 2), Web of Knowledge, GreenFILE (Ebsco), ScienceDirect, ProQuest Dissertations and Google Scholar. The literature encompasses books, papers and doctoral dissertations about complexity, chaos and their application in building construction, as well as entropy algorithms used for complexity measurement in different fields. Most of the conceptual complexity research and complexity measurement papers are published in top ranked journals of complexity research and chaos theory: *Chaos and Complexity, Chaos, Solutions and Fractals* and *Physical Review*. The complexity research papers related to the construction industry are all published in high-impact journals in building construction, such as *International Journal of Project Management, Construction Management and Economics, Construction Engineering and Management, Automation in Construction,* and *Architectural Engineering and Design Management*. These journals are the major route to track current research in complexity, especially complexity research in building construction. Papers from these target journals also provide solid background and introduction of commonly used entropy algorithms for complexity measurement to be further compared and analyzed.

Based on the literature review, the next phase of study uses conceptual analysis and metonymic mapping to summarize the complexity disciplines, to build a theoretical frame that is applicable and found in construction, and to select potential entropy models based on the theoretical frame. First, the study summarizes common definitions of complex system and complexity in building construction, lists the characteristic of complex systems, and then applies them to construction performance. Next, it clarifies the relationship between complexity and chaos, and verifies it through graphic findings from the construction industry. This forms the foundation for a theoretical framework of complexity in building construction from three levels – complex system, complex theory and science of complexity. From the perspective of complexity science, the four most commonly used and the most well-known entropy algorithms, listed in Table 7, are selected as the potential complexity measurements for further comparison.

Table 7 Summary of Entropy Models for Complexity or Chaos

| **Entropy Model** | **Definition** | **Author & Year** |
|---|---|---|
| Approximate Entropy (ApEn) | ApEn is a technique used to quantify the amount of regularity and the unpredictability of fluctuations over time-series data. | Pincus (1991)<br>(Has been cited 2859 times). |
| Sample Entropy (SampEn) | It is a modification of ApEn. It used extensively for assessing the complexity of a physiological time-series signal. But it does not include self-similar patterns as ApEn does. | Grassberger (1988) (Has been cited 505 times); Richman and Moorman (2000) (Has been cited 2184 times). |
| Permutation Entropy (PEn) | PerEn is an appropriate complexity in the presence of dynamical and observational noise. | Bandt and Pompe (2002) Has been cited 838 times). |

### *3.3 Pilot Comparative Study*

The comparative method is one of the oldest research methods in the social sciences. The objective of using a pilot comparative study as the second phase of this research is to conduct initial analysis for the four proposed entropy algorithms, along with Lyapunov Exponent and Six Sigma methods. In the pilot comparison stage, three (3) different entropy models of complexity measures are compared  Approximate Entropy (ApEn), Sample Entropy (SampEn) and Permutation Entropy (PEn). They are also compared with other commonly used complexity measurement methods--Lyapunov Exponent, and the method for building construction performance evaluation known as Sig Sigma Method.

Not only is a qualitative comparison for the proposed methods conducted based on the literature and definitions, a quantitative comparison is also made based on ten (10) pilot cases. These ten (10) pilot cases were randomly picked from the twenty-two (22)  real construction cases that Fernandez-Solis used for his study of chaos theories in building construction (Fernandez-Solis, 2013). All these cases used Percent Plan Complete (PPC, also known as Promise Plan Complete) to represent building construction performance. According to Lean Construction Institution (2015), PPC is a basic measure of how well the planned building construction system is working - calculated as the "number of assignments completed on the day specified" divided by the "total number of assignments made for the week." It measures the percentage of assignments that are

100% complete as planned. This represents the overall reliability of production planning and that of workflow, rather than focusing only on a certain perspective, such as schedule, cost and safety (Ballard, 1999). The following Table 8 and Figure 7 list the PPC report for each pilot case.

Table 8 PPC Reports for Pilot Cases

| Week | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 | Case 6 | Case 7 | Case 8 | Case 9 | Case 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 16.67% | 51.98% | 19.00% | 40.96% | 40.15% | 65.19% | 60.85% | 36.88% | 32.50% | 53.57% |
| 2 | 15.63% | 47.22% | 45.83% | 32.58% | 30.88% | 45.92% | 42.86% | 29.91% | 25.00% | 42.62% |
| 3 | 55.83% | 41.27% | 66.11% | 44.55% | 39.52% | 81.51% | 76.07% | 40.97% | 31.99% | 58.28% |
| 4 | 55.83% | 50.00% | 47.22% | 67.82% | 69.12% | 95.00% | 93.33% | 81.85% | 55.95% | 88.71% |
| 5 | 78.93% | 51.67% | 82.90% | 61.79% | 52.94% | 79.72% | 74.40% | 67.50% | 42.86% | 80.82% |
| 6 | 70.89% | 23.33% | 69.81% | 63.79% | 88.24% | 95.00% | 92.59% | 52.08% | 71.43% | 83.44% |
| 7 | 61.61% | 61.11% | 80.11% | 64.88% | 72.06% | 95.00% | 89.68% | 50.00% | 58.33% | 84.86% |
| 8 | 49.17% | 16.67% | 46.11% | 57.59% | 59.04% | 64.29% | 60.00% | 64.11% | 47.80% | 75.32% |
| 9 | 95.00% | 95.00% | 73.48% | 76.45% | 88.79% | 84.69% | 79.05% | 71.25% | 71.88% | 95.00% |
| 10 | 60.28% | 60.28% | 78.94% | 75.98% | 95.00% | 84.82% | 79.17% | 77.08% | 80.95% | 99.38% |

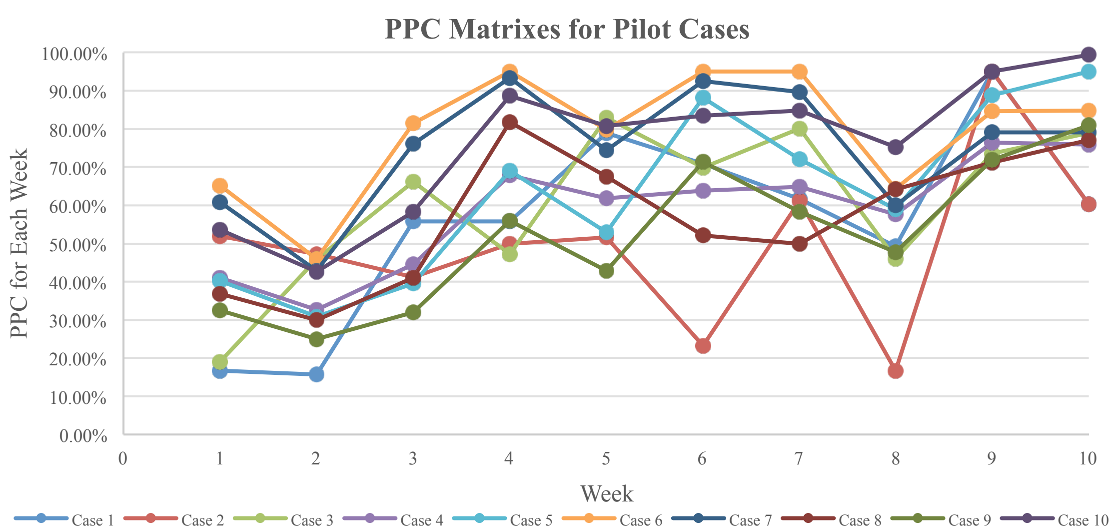



Figure 7 PPC Matrixes for Pilot Cases

For these pilot cases, the measured results based on each different model are compared from three different perspectives: the degree of calculation difficulty, sensitivity and dependence. On the basis of both the qualitative comparison and quantitative comparison of pilot cases, the criteria for further comparison and preliminary conclusions are summarized.

### *3.4 Simulations and Comparative Analysis*

In order to test the proposed entropy algorithms in the different construction performance scenarios and to make sure the proposed quantitative measuring model could be accepted in real construction work, Monte Carlo simulation is used to generate more random numbers for each pilot case. This mimics real construction situations over long periods as case studies, which is an ideal methodology when a holistic, in-depth investigation is needed (Feagin et al., 1991). Based on the simulated cases, the entropy algorithms are further compared, the complexity of construction is measured and also some different circumstances are tracked.

At this stage, two rounds of simulation are conducted with MATLAB to generate random numbers based on the pilot cases for the purpose of representing different circumstances in building construction performance as illustrated in Figure 8.

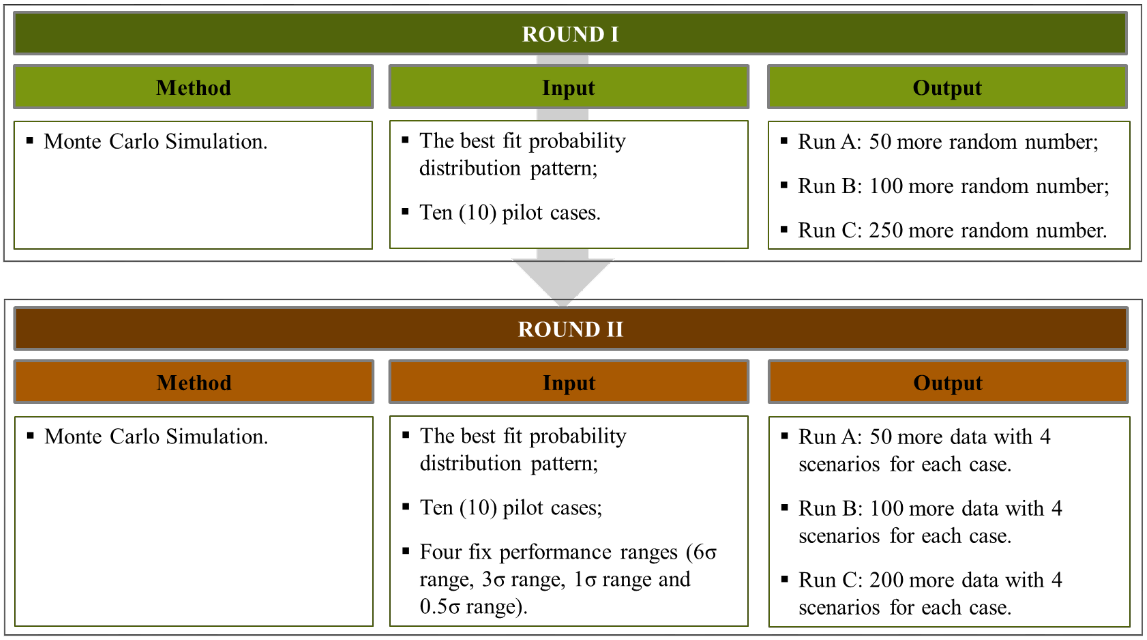



Figure 8 Process for Two-Round Simulation based on Pilot Cases

The first round simulation is aimed to generate more random numbers following the same pattern of each pilot case to test the workability of proposed entropy models in real long term complex building construction systems. For each pilot case, twenty-two (22) different probability distribution scenarios are tested in order to find the specific pattern with the highest fitness to generate random numbers. The probability distributions tested for each pilot case include Beta, Binomial, Birnbaum-Saunders, Burr, Exponential, Extreme Value, Gamma, Generalized Extreme Value, Generalized Pareto, Inverse Gaussian, Log-Logistic, Logistic, Lognormal, Nakagami, Negative Binomial, Non-parametric, Normal, Poisson, Rayleigh, Rician, t Location-Scale, and Weilbull. According to the probability distribution pattern with the highest fitness, 50, 100 and 250 additional numbers are generated for further comparative analysis of proposed entropy algorithms.

The second round of simulation mimics different circumstances in building construction systems, rather than simply increasing the sample size. It generates more random numbers for each pilot case within a certain range by using MATLAB. In the round two simulation, three runs are conducted, like in the previous simulation: Run II-A with 50 more numbers, Run II-B with 100 more numbers and Run II-C with 200 more numbers. However, in each run, the measurement results are not compared across different cases, they are compared based on different performance patterns for the same case. For each pilot case, four different simulated scenarios are generated. Simulated random numbers fall in the range of (Mean-3*Standard Deviation, Mean-3*Standard Deviation) for scenario a of each case, (Mean-1.5*Standard Deviation, Mean-1.5*Standard Deviation) for scenario b, (Mean-0.5*Standard Deviation, Mean-0.5*Standard Deviation) for scenario c and (Mean-0.25*Standard Deviation, Mean-0.25*Standard Deviation) for the last scenario, which is scenario d. Selected entropy models are applied for complexity calculation of each simulated circumstance set by the researcher to check the performance of the selected model.

Conclusions and recommendations are based on the results of all the simulated circumstances, verifying the effectiveness of selected entropy models, and also showcasing how to use the model in real construction work.

# CHAPTER IV

# PROPOSED ENTROPY ALGORITHMS

Various measures of complexity have been developed to compare time series and distinguish regular (e.g., periodic), chaotic, and random behavior. The main types of complexity parameters are entropies, fractal dimensions, and Lyapunov exponents. For the entropy algorithm, there are also several perspectives to measure complexity. Different entropy models have been developed or modified to accommodate the characteristics of data source, project type and research requirements. From the perspective of chaos theory, as summarized in Chapter II Literature Review, the following comparison will focus on three commonly used entropy algorithms, a method cited more than 500 times by different authors. They are Approximate Entropy (ApEn), Sample Entropy (SampEn) and Permutation Entropy (PEn).

## *4.1 Approximate Entropy (ApEn)*

Approximate Entropy (ApEn) was introduced as a quantification of regularity and complexity in noisy, short time-series data by Pincus in 1991 in “Approximate entropy as a measure of system complexity” published in *Proceedings of the National Academy of Sciences* (Pincus, 1991). This is the original appearance of the ApEn algorithm and the paper has been cited 2827 times until now. Mathematically, ApEn represents the rate of entropy for an approximating Markov chain to a process (Pincus, 1991), which appears to have potential application to a wide variety of relatively short (greater than 100 points) and noisy time-series data. A low value for entropy indicates that the time series is deterministic; a high value indicates randomness.

The length of compared runs, which is represented as m, and effective filtering level, which is represented as r, are two input parameters that must be fixed for computing ApEn. Given N data points {u(i)}, form vector sequences x(l) through x(N-m+1), defined by x(i)= [u(i), ... , u(i+m-1)]. These vectors represent m consecutive u values, commencing with the ith point. Define the distance d[x(i), x(j)] between vectors x(i) and x(j) as the maximum difference in their respective scalar components. Use the sequence x(l), x(2), ... , x(N-m+1) to construct, for each I ≤ N-m +1, $C_i^m(r)$ = (number of j≤N-m+1 such that d[x(i), x(j)] ≤r)/(N-m+1). The $C_i^m(r)$'s measure, within a tolerance r, the regularity, or frequency, of patterns similar to a given pattern of window length m. Define $\phi^m(r) = (N + m - 1)^{-1} \sum_{i=1}^{N-m+1} ln\, C_i^m(r)$, where ln is the natural logarithm, then define the parameter Approximate Entropy ApEn(m, r)=$lim_{N\to\infty}[\phi^m(r) - \phi^{m+1}(r)]$. It measures the likelihood that pattern runs that are close for m observations will remain close on the next incremental comparisons.

The procedure of ApEn algorithm, summarized in the following steps, was coded in MATLAB (see Appendix A):

1. Form a time series of data {u(1), u(2), …, u(N)}. There are N raw data values as measured at equally spaced times.
2. Fix length of compared runs (m) and effective filtering level (r). Typically choose m=1 or m=2, and r depends greatly on the application.
3. Form a sequence of vectors x(1), x(2), …, x(N-m+1), in $R^m$, real m-dimensional space defined by x(i)= [u(i),u(i+1), ... , u(i+m-1)].
4. Use the sequence x(1), x(2), …, x(N-m+1) to construct, for each i, $1 \leq i \leq N + m - 1$:

$$C_i^m(r) = \frac{(\text{number of x(j) such that d[x(i), x(j)]} \ \leq \text{r})}{(\text{N} - \text{m} + 1)}$$

   in which d[x, x*] is defined as $d[x, x^*] = \max\ |\text{u(a)} - u^*(a)|$.

   The u(a) are the m scalar components of x. d represents the distance between the vectors x(i) and x(j), given by the maximum difference in their respective scalar components. Note the j takes on all values, so the match provided when i = j will be counted (the subsequence is matched against itself).
5. Define

$$\phi^m(r) = (N + m - 1)^{-1} \sum_{i=1}^{N-m+1} ln\,(C_i^m(r)).$$

6. Calculate Approximate Entropy ApEn as

$$\text{ApEn(m, r)}=lim_{N\to\infty}[\phi^m(r) - \phi^{m+1}(r)].$$

### *4.2 Sample Entropy (SampEn)*

Using the ApEn algorithm requires each template to contribute a defined nonzero probability. This is secured by permitting self-matching of each template. However, this makes ApEn a biased estimate, which indicates more similarity than is truly present for finite N (Pincus & Goldberger, 1994). An estimate of a statistic is biased if its expected value is not equal to the parameter it estimates.

Sample entropy (SampEn) was developed to reduce the bias caused by self-matching. The idea to associate sample correction with entropy estimation, Sample Entropy (SampEn), comes from the work of Grassberger (1988). The name refers to the applicability to time series data sampled from a continuous process. Sample entropy, a modification of approximate entropy, is used extensively for assessing the complexity of a physiological time-series signal, thereby diagnosing a diseased state. It examines a time series for similar epochs and assigns a non-negative number to the sequence, with large values corresponding to more irregularity in the data, and more complexity in the system (Richman & Moorman, 2000). Richman and Moorman (2000) wrote "Physiological time-series analysis using approximate entropy and sample entropy," which has become the most widely cited paper for Sample Entropy Algorithms (cited 2168 times to present).

SampEn differs from ApEn in two ways. First, SampEn does not count self-matches. This is justified since entropy is a measure of the rate of information production, and in this context, comparing data with themselves is meaningless. Second, SampEn does not use a template-wise approach when estimating conditional probabilities. It only requires that one template find a match of length m+1, then it computes the logarithm of a probability associated with the time series as a whole (Al-Angari & Sahakian, 2007).

The length of compared runs, which is represented as m, and effective filtering level, which is represented as r, are two input parameters that must be specified to compute ApEn. Given N the length of time series $\{u(i)\}$, Sample Entropy (SampEn) is 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. Thus, a lower value of SampEn also indicates more self-similarity in the time series (Richman & Moorman, 2000).

Formally, given N data points from a time series $\{u(1), u(2), \ldots, u(N)\}$, SampEn, calculated as follows (Richman & Moorman, 2000), was coded in MATLAB (see Appendix B):

1. Fix length of compared runs (m) and effective filtering level (r). Although m and r are critical in determining the outcome of SampEn, no guidelines exist for optimizing their values. In principle, the accuracy and confidence of the entropy

estimate improve as the number of length m matches increases. The number of matches can be increased by choosing small m (short templates) and large r (wide tolerance). Typically, one chooses m=2, and r between 0.1 and 1 times the standard deviation of the original time series {u(i)}, as recommended by Pincus (2001).

2. Form vector sequences of size m, $x_m(1)$, $x_m(2)$, …, $x_m(N-m+1)$ in $R^m$, real m-dimensional space defined by $x_m(i)= [u(i),u(i+1), \dots , u(i+m-1)]$. These vectors represent m consecutive x values starting with the $i^{th}$ point.
3. Define the distance between vectors $x_m(i)$ and $x_m(j)$, $d[x_m(i), x_m(j)]$, as the absolute maximum difference between their scalar components:

$$d[x_m(i), x_m(j)] = max_{k=0,\dots,m-1}(|u(i+k) - u(j+k)|)$$

4. For a give x(i), count the number of j ($1 \leq j \leq N - m, j \neq i$), denoted as Bi, such that the distance between $x_m(i)$ and $x_m(j)$ is less than or equal to r. Then, for $1 \leq i \leq N - m$:

$$B_i^m(r) = \frac{1}{(N - m - 1)} B_i$$

5. Define $B^m(r)$ as:

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r)$$

6. Increase the dimension to m + 1 and calculate $A_i$ as the number of $x_{m+1}(i)$ within r of $x_{m+1}(j)$, where j ranges from 1 to N – m ($i \neq j$). Then, $A^m(r)$ is defined as:

$$A_i^m(r) = \frac{1}{(N - m - 1)} A_i$$

7. Set $A^m(r)$ as:

$$A^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} A_i^m(r)$$

Thus, $B^m(r)$ is the probability that two sequences will match for m points, whereas $A^m(r)$ is the probability that two sequences will match for m + 1 points.

8. Finally, SampEn can be defined as:

$$SampEn(m,r) = \lim_{N\to\infty}\{-\ln[\frac{A^m(r)}{B^m(r)}]\}$$

which is estimated by the statistic

$$SampEn(m,r,N) = -\ln[\frac{A^m(r)}{B^m(r)}]$$

### *4.3 Permutation Entropy (PEn)*

Permutation entropy was introduced in 2002 by Bandt and Pompe as a complexity measure for time series data based on a comparison of neighboring values. Their paper, "Permutation entropy: a natural complexity measure for time series," has been cited more than 806 times until now.

Roughly speaking, permutation entropy replaces the probabilities of length-L symbol blocks in the definition of Shannon entropy by the probabilities of length-L ordinal patterns, each pattern being a digest of the ups and downs of L consecutive elements of a time series. Since its introduction, the permutation entropy algorithm, along with different tools based on ordinal patterns, has found a number of interesting applications.

When quantifying complexity for a given time series {u(i)} that length is equal to T,, permutation entropy reflects the rank order of successive

$u_i$ in sequences of length n and thus is defined as:

$$H(n) = -\sum_{j=1}^{n!} p(\pi)\log\left(p(\pi)\right)$$

The $p(\pi)$ represent the relative frequencies of the possible patterns of symbol sequences, termed permutations (see Fig. 1). This relative frequency is determined by

$$p(\pi) = \frac{\#\{t|t \le T-n, (u_{t+1}, \ldots, u_{t+n})\ has\ type\ \pi\}}{T-n+1}$$

The permutation entropy per symbol is given by

$$h(n) = -\frac{1}{n-1}\sum_{j=1}^{n!} p(\pi)\log\left(p(\pi)\right)$$

in order to compare permutation entropies with different n. Besides the normalization using n – 1, there are also other approaches (e.g., normalization with $\log_2(n!)$ to get $0 \le h_n \le 1$).

PEn essentially measures information based on the occurrence or absence of certain permutation patterns of the ranks of values in a time series. To compute the PEn for a given time series {u(i)} of length T, following are seven steps listed in detail and in Appendix C (Bandt & Pompe, 2002):

1. Define the order of permutation n. That leads to the possible permutation patterns $\pi_j$ (j = 1, .., n!) which are built from the numbers 1, ..., n.
2. Initialize i = 1 as the index of the considered time series $\{u(i)\}_{i=1,...,T}$ and the counter $z_j = 0$ for each $\pi_j$ .
3. Calculate the ranks of the values in the sequence $u_i, ..., u_{i+n-1}$ which leads to the rank sequence $r_i, ..., r_{i+n-1}$. The ranks are the indices of the values in ascending sorted order.
4. Compare the rank sequence of step 3 with all permutations patterns and increase the counter of the equal pattern $\pi_k = r_i, ..., r_{i+n-1}$ by one ($z_k = z_k + 1$).
5. If $i \leq T - n$ then increase i by one ($i = i + 1$) and start from step 3 again. If $i > T - n$ go to the next step.
6. Calculate the relative frequency of all permutations $\pi_j$ by means of $p(\pi)_j = \frac{z_j}{\sum z_k}$ as an estimation of their probability $p_j$ .
7. Select all values of $p(\pi)_j$ greater than 0 (since empty symbol classes yield 0log0 = 0) and calculate the permutation entropy by

$$H(n) = -\sum_{j=1}^{n!} p(\pi)\log\,(p(\pi))$$

$$h(n) = -\frac{1}{n-1}\sum_{j=1}^{n!} p(\pi)\log\,(p(\pi))$$

# CHAPTER V

# PILOT COMPARATIVE STUDY AND RESULT

For the pilot comparative study of proposed entropy algorithms, ten (10) cases were randomly selected from the twenty-two (22) real construction cases that Fernandez-Solis' used for his study of chaos theories in building construction (Fernandez-Solis, 2013). All ten cases used Percent Plan Complete (PPC) to represent building construction performance. The original PPC reports for the selected pilot cases are presented in Table 8 and Figure 7 in Chapter III. These cases will be used for preliminary tests of the three (3) proposed Entropy Algorithms. The comparison of complexity measurement results from all the models, along with Lyapunov Exponent and Six Sigma, could provide not only initial results for further testing, but also some critical criteria for further simulation and comparative analysis.

## *5.1 Six Sigma Analysis of Pilot Cases*

The Six Sigma approach was initiated in the early 1980s at Motorola, but it gained little attention until General Electric and AlliedSignal successfully adopted it in their organizations in the late 1990s. Because of its practical successes, the study and development of the Six Sigma method have been widely conducted in different fields. Numerous strategies, tools, techniques and principles have been invented; there are also plenty of different definitions of Six Sigma, each with its own point of emphasis. The fundamental definition is provided by Harry and Schroeder (2000), who were the principal developers of the Six Sigma program at Motorola. They define Six Sigma as "a disciplined method of using extremely rigorous data gathering and statistical analysis to pinpoint sources of errors and ways of eliminating them." On the basis of this definition, Six Sigma have been further developed from numerous perspectives, such as a strategic approach (Snee, 2000), a statistical measure and management philosophy (Chowdhury, 2001; Pande et al., 2002), and a comprehensive and flexible system for achieving and maximizing business success (Pande et al., 2000). Among all these definitions, this study considers Six Sigma to be a statistical measure used to measure the performance of processes against anticipated plan, which is also known as a "technical" definition of Six Sigma compared to the business point of view (Kwak & Anbari, 2006).

The statistical Six Sigma principle is a universal quality metric that measures the performance of specified process, which in this study is the building construction process. High sigma values indicate better processes with fewer numbers of defects, which would prevent the realization of expectations. A simplified Six Sigma Conversion Table that could be used in construction was summarized by Pheng and Hui (2004), as shown in Table 9.

Table 9 Simplified Sigma Conversion Table (Adapted from Pheng & Hui, 2004)

| Yield=Percentage of items achieve the anticipated plan | Defects per million Opportunities (DPMO) | Sigma Level |
|---|---|---|
| 30.9 | 690,000 | 1 |
| 69.2 | 308,000 | 2 |
| 93.3 | 66,800 | 3 |
| 99.4 | 6,210 | 4 |
| 99.98 | 320 | 5 |
| 99.9997 | 3.4 | 6 |

As there is no proven statistical rule for using Six Sigma analysis in the construction field, two different rules were set up in this study for the pilot comparison of ten cases in order to find their sigma levels and for further testing: I) if the Weekly PPC was larger than the mean of each case minus its standard deviation and less than the mean plus standard deviation, it would be considered as achieving the anticipated plan; II) if the Weekly PPC was larger than the average PPC of each case, it would be considered as achieving the anticipated plan. Figures 9 and 10 showcase the Six Sigma analysis for Case 1 based on different rules. These two rules were tested in the pilot cases, and one of them was selected for the study.

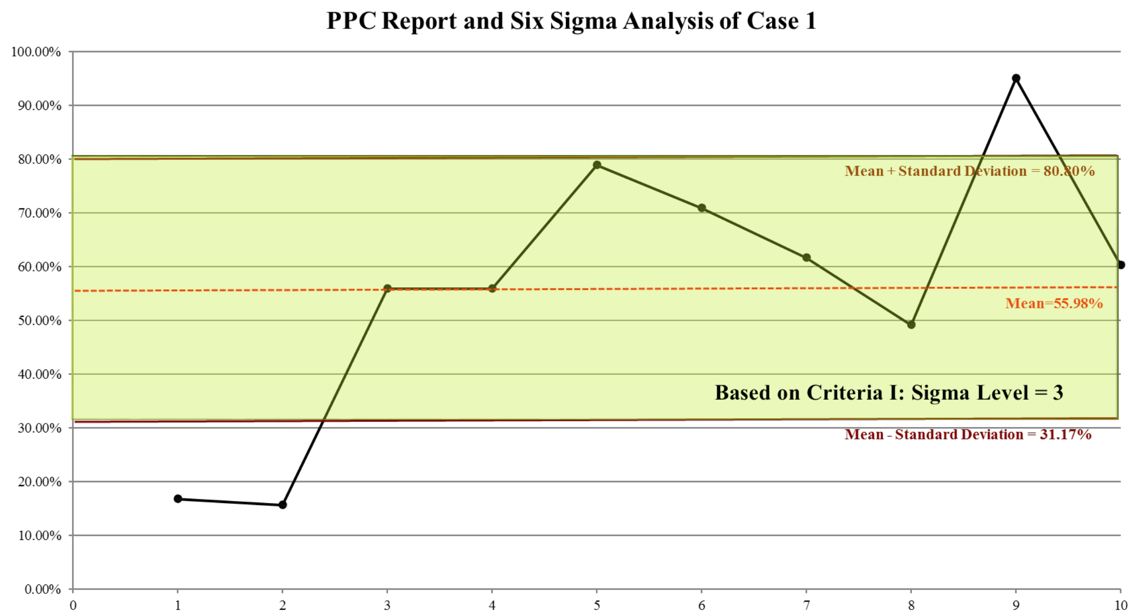



Figure 9 Six Sigma Analysis for Case 1 based on Criterion I

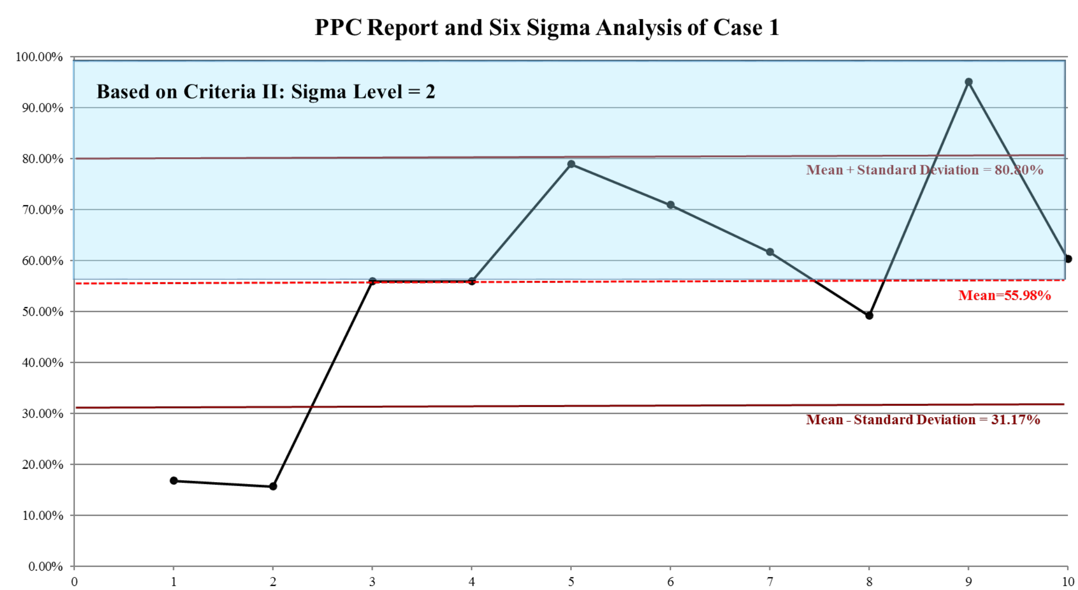



Figure 10 Six Sigma Analysis for Case 1 based on Criterion II

Following the analysis of Case 1, the same Six Sigma Analysis was conducted on the rest of the pilot cases. According to these two criteria, the sigma levels of each case are listed in the Table 10.

Table 10 Six Sigma Analysis for Pilot Cases

| **CASE** | **Mean** | **Standard Deviation** | **Criteria I** | | **Criteria II** | |
|---|---|---|---|---|---|---|
| | | | **Yield** | **Sigma Level** | **Yield** | **Sigma Level** |
| **1** | 55.98% | 24.82% | 70.00% | 3 | 50.00% | 2 |
| **2** | 49.85% | 21.52% | 70.00% | 3 | 60.00% | 2 |
| **3** | 60.95% | 20.63% | 80.00% | 3 | 60.00% | 2 |
| **4** | 58.64% | 14.79% | 60.00% | 2 | 60.00% | 2 |
| **5** | 63.57% | 22.76% | 40.00% | 2 | 50.00% | 2 |
| **6** | 79.11% | 16.13% | 90.00% | 3 | 70.00% | 3 |
| **7** | 74.80% | 16.17% | 70.00% | 3 | 60.00% | 2 |
| **8** | 57.16% | 17.80% | 60.00% | 2 | 50.00% | 2 |
| **9** | 51.87% | 19.08% | 40.00% | 2 | 50.00% | 2 |
| **10** | 76.20% | 18.73% | 60.00% | 2 | 60.00% | 2 |

According to Table 10, sigma levels for the ten pilot cases, based on both criteria, only offer a rough ranking for the performance level. Within each level, the performance evaluations for different cases are not significant. As presented in Figure 11, according to Six Sigma analysis of criterion I, Cases 1, 2, 3, 6, and 7 have the same Sigma Level. However, significantly different patterns for these cases are easily observed from their PPC metrics. Even the cases with the same yield percent (Cases 1, 2, and 7) take on different trends. As a result, it could be preliminarily concluded that Six Sigma as a statistical approach only picks up and ranks the projects for a wide range rather than reflecting fluctuation within this range. That is to say, the Six Sigma Statistical approach is not a sensitive enough measurement for building construction performance and its complexity.

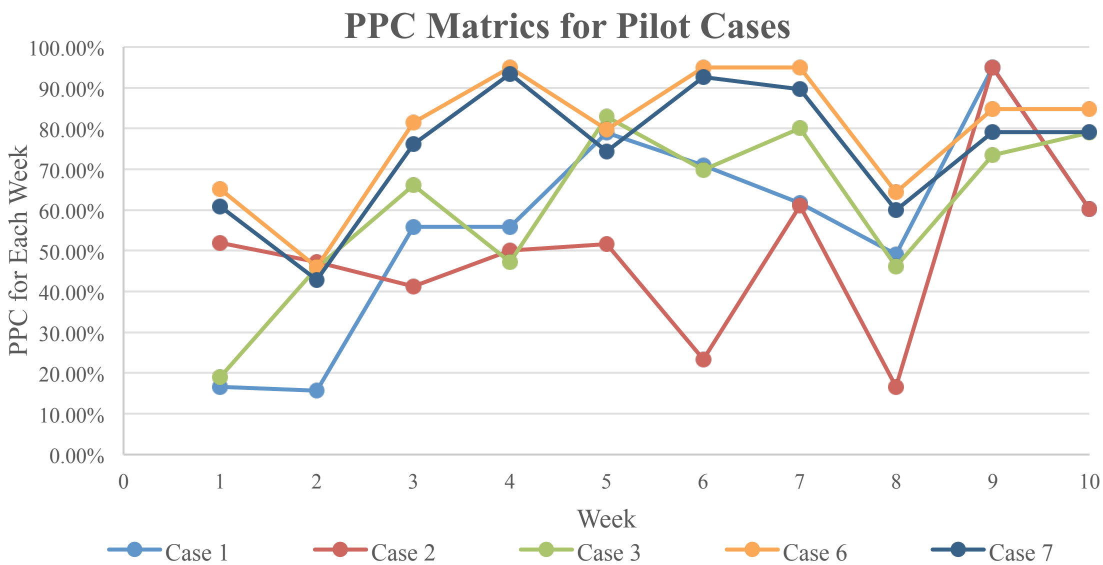



Figure 11 PPC Metrics of Selected Cases with Same Sigma Level.

From the ranked results based on the two predetermined criteria as shown in Table 10, the results based on criterion II were almost the same. The same statistical phenomenon becomes more obvious when the sample size increases. As a result, this criterion will be eliminated from the pilot study. Thus, the first criterion of Six Sigma Analysis was chosen for the following work.

## *5.2 Complexity Measurement of Pilot Cases*

To test the selected entropy models, each of the three entropy algorithms were used to measure the complexity of the ten pilot cases in turn. From the definitions and algorithm procedures discussed in Chapter IV, we specify the input time-series data source, which is each individual pilot case; each entropy model also contains several parameters required for the calculation. With different parameter set-ups, the measurement results should be different. However, there are no clearly defined selection principles for each parameter in the current research field. The pilot case study will also aim to test different parameter combinations for each proposed entropy algorithm.

Approximate Entropy is first tested using the pilot cases, for which fixed length of compared runs (m) and effective filtering level (r) are the parameters required. The parameter combination tested in the pilot cases and the measurement results are listed in Table 11. Because it shares similar principles of calculation with Approximation Entropy, the idea and parameters required for Sample Entropy are the same, whose results are shown in Table 12. Permutation Entropy, measures complexity based on the pattern of the metric. For this reason, only the length of the compared unit is required for calculation. Its parameter test and calculation results are listed in Table 13.

Table 11 Complexity Measurement for Pilot Cases Using ApEn

| ApEn | CASE 1 | CASE 2 | CASE 3 | CASE 4 | CASE 5 | CASE 6 | CASE 7 | CASE 8 | CASE 9 | CASE 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **m=1, r=0.1*Std** | 0.3105 | 0.3629 | 0.3629 | 0.1719 | 0.1719 | 0.5015 | 0.3105 | 0 | 0.1719 | 0.0333 |
| **m=1, r=0.2*Std** | 0.4609 | 0.4609 | 0.6483 | 0.2376 | 0.3105 | 0.5672 | 0.5554 | 0.1719 | 0.3105 | 0.0000 |
| **m=1, r=0.5*Std** | 0.7338 | 0.4529 | 0.5825 | 0.3159 | 0.4501 | 0.7633 | 0.7633 | 0.6653 | 0.4501 | 0.3916 |
| **m=1, r=Std** | 0.3943 | 0.5075 | 0.4627 | 0.2537 | 0.6326 | 0.4129 | 0.4047 | 0.4406 | 0.6326 | 0.2537 |

*Std = Standard Deviation of each case

Table 12 Complexity Measurement for Pilot Cases Using SampEn

| SampEn | CASE 1 | CASE 2 | CASE 3 | CASE 4 | CASE 5 | CASE 6 | CASE 7 | CASE 8 | CASE 9 | CASE 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **m=2, r=0.1*Std** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| **m=2, r=0.2*Std** | NaN | NaN | NaN | Inf | NaN | Inf | Inf | NaN | NaN | Inf |
| **m=2, r=0.5*Std** | Inf | 0.8473 | 0.6931 | 0.6931 | 1.0986 | Inf | Inf | NaN | 1.0986 | 0.9163 |
| **m=2, r=Std** | 0.3567 | 0.5878 | 1.0986 | 0.3185 | 1.2528 | 0.2877 | 0.2231 | 0.4700 | 1.2528 | 0.3185 |

*Std = Standard Deviation of each case

Table 13 Complexity Measurement for Pilot Cases Using PEn

| PEn | CASE 1 | CASE 2 | CASE 3 | CASE 4 | CASE 5 | CASE 6 | CASE 7 | CASE 8 | CASE 9 | CASE 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **n=2** | 0.6365 | 0.6870 | 0.6365 | 0.6870 | 0.6870 | 0.6365 | 0.6870 | 0.6870 | 0.6870 | 0.6365 |
| **n=3** | 1.0822 | 1.4942 | 1.5596 | 1.5596 | 1.5596 | 1.3209 | 1.7329 | 1.3209 | 1.5596 | 1.4942 |
| **n=4** | 1.7479 | 1.7479 | 1.9459 | 1.9459 | 1.9459 | 1.7479 | 1.9459 | 1.7479 | 1.9459 | 1.7479 |

### *5.3 Lyapunov Exponent for Pilot Cases*

In current research about chaos and complexity, the Lyapunov exponent is also used to produce maps of complexity and to represent the complex levels of time series data. The algorithm chosen for complexity measurement on the basis of sensitivity to initial conditions of the specific dynamical system is called the Lyapunov exponent method

(Wolf et al., 1985). It provides a quantitative measure of sensitivity to initial conditions and is one of the most useful dynamical diagnostics for chaotic and complex systems. Based on a summary by Zeng, Eykholt and Peilke (1991), any system that contains at least one positive Lyapunov exponent can be defined as chaotic and complex.

Because it is a dynamical and complex diagnostic developed and used before entropy algorithms, there are plenty of methods for calculating the Lyanupov Exponent. For this study, I apply the Maximal Lyapunov Exponent (or so called Largest Lyapunov Exponent) as the complex prediction and measurement method to compare with the proposed entropy algorithms (Kantz, 1994). Compared to the tradition Lyapunov Exponent calculation, this algorithm makes use of the statistical properties of the local divergence rates of nearby trajectories. As such, it does not depend on knowledge of the correct embedding dimension or on other parameters. The calculation procedures in order to find the Maximum Lyapunov were coded in MATLAB (see Appendix D). The Maximal Lyapunov Exponent for each pilot case is found in Table 14.

Table 14 Maximal Lyapunov Exponent for Pilot Cases

| Case | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Maximal Lyapunov Exponent | 0.3466 | 0.6838 | 0.5959 | 0.1238 | 0.7115 | 0.2101 | 0.1349 | 0.3475 | 0.7115 | 0.1201 |

### *5.4 Descriptive Analysis of Pilot Cases Study*

From the Six Sigma analysis, Sigma Levels for the ten pilot cases based on both criteria only offer a rough rank for performance level. Within each level, the performance evaluations for different cases are not significant. However, significantly different patterns for these cases are easily observed from their PPC metrics. As a result, it could be preliminarily concluded that Six Sigma analysis only picks up and ranks the projects across a wide range, rather than reflecting fluctuation within this range, which means the measured results are not sensitive to the pattern and fluctuation of the case. That is to say, Six Sigma Statistical approach is not a sensitive enough measurement for analyzing building construction performance and its complexity. However, even though it has low sensitivity for complexity measurement, the rough ranking results from Six Sigma Analysis may still be used as a predetermined verification condition to test the proposed complexity measurement algorithms, as it has been proven and utilized in previous construction performance analyses.

Both Approximate Entropy and Sample Entropy capture the dynamical level of performance or complex system from the perspective of information contained in each case. Their calculation procedures and required parameters also share some common principles. However, from their measurement results, ApEn works better for a limited amount of data points (only 10 numbers for each case), whereas SampEn does not provide precise results for a small number of comparisons and with less tolerance. In addition, for ApEn, even the measured complexity results are sensitive to different cases (different number series), but there is still no unified rank based on the different input combinations, which are the compared template units and tolerance. This means that further tests with simulated cases with greater captured numbers are necessary.

From Table 11, we see that all parameter combinations of ApEn are workable with the small numbers. For this reason, to get more sensitive results in the large sample size complexity measurement in the following analysis, we introduce a higher parameter criterion, which means a smaller tolerance (r=0.1*Std or r=0.2*Std). Sample Entropy (see Table 12) does not work well with a limited number of cases. In order to compare the complexity level among the cases and increase the sensitivity of this method, a small sample size requires that we use a wider tolerance. In addition, the workability of Sample Entropy, especially for the smaller tolerances, should also be further tested using a larger sample size.

The Permutation Entropy algorithm measures entropy from a different perspective compared to ApEn and SampEn. It focuses on the pattern of the time-series data, rather than on the data itself. For example, the two matrices [4, 2, 8] and [2, 1, 4] have the same pattern but with different numbers. Based on this idea, the results of PEn show no different complexity levels for data with the same pattern of fluctuation. That is to say, for different time series, compared with Six Sigma Analysis, PEn is sensitive to fluctuation patterns, but not to data values.

The Maximal Lyanupov Exponent method is sensitive to both the pattern and value of time-series data. However, the differences among all the measurement methods are not as significant as with Sample Entropy, but they are still more significant than with Approximate Entropy.

From the above discussion, some inferential conclusions are made:

1. Compared with other complex performance measurement methods, Sample Entropy does not work well with small sample size;
2. The level of sensitivity for each methods:
   Six Sigma Analysis < PEn < SampEn = ApEn = Maximal Lyanupov Exponent.
3. The significant levels among the measured complexity results for each methods:
   ApEn < Maximal Lyanupov Exponent < SampEn.

In order to better understand the proposed entropy algorithms, we also checked and compared their results. Table 15 ranks the pilot cases from the most complex one (the hardest to be controlled) to the least complex one (could be easily controlled) using different models. Figure 12 plots the complexity measurement results based on different algorithms in a single chart. From both the figure and the table, we see that ranks are different based on different methods, as they are from different perspectives. The ranks based on PEn with different compared unit numbers are quite similar. In addition, the rank based on Sample Entropy and the one based on Maximal Lyanupov Exponent are roughly the same (although not exactly same), which is worthy of further testing.

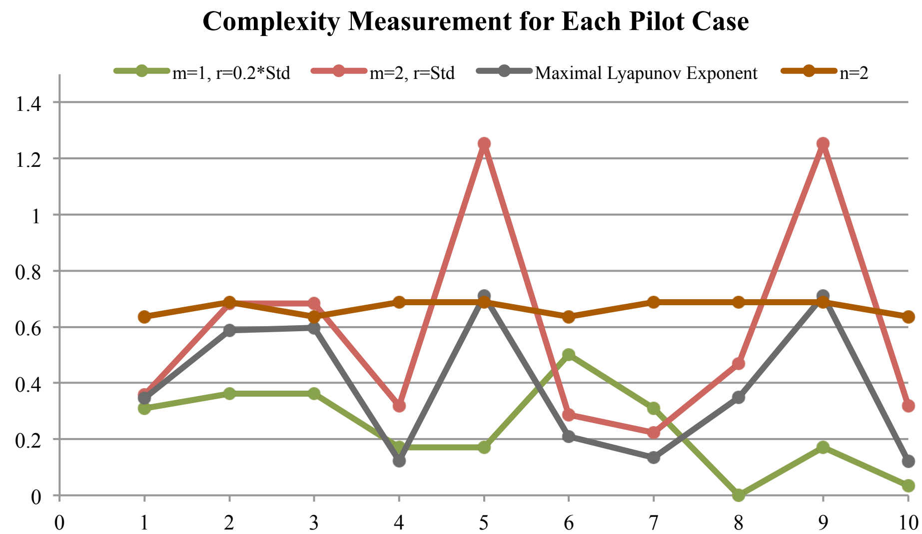



Figure 12 Complexity Measurements for Each Pilot Case

Table 15 Pilot Cases Rank by Proposed Methods

| Method | Cases Rank from the Most Complex to the Least | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Six Sigma** | 9 | 5 | 8 | 4 | 10 | 7 | 1 | 2 | 3 | 6 |
| **ApEn(m=1,r=0.5*Std)** | 3 | 10 | 6 | 1 | 5 | 8 | 4 | 7 | 2 | 9 |
| **ApEn(m=1,r=Std)** | 3 | 4 | 1 | 9 | 8 | 2 | 7 | 6 | 5 | 10 |
| **SampEn(m=2, r=Std)** | 5 | 9 | 3 | 2 | 8 | 1 | 4 | 10 | 6 | 7 |
| **PEn (n=2)** | 7 | 4 | 5 | 9 | 8 | 2 | 3 | 10 | 6 | 1 |
| **PEn (n=3)** | 7 | 4 | 5 | 9 | 3 | 2 | 10 | 8 | 6 | 1 |
| **PEn (n=4)** | 7 | 4 | 5 | 9 | 3 | 2 | 10 | 8 | 6 | 1 |
| **Maximal Lyapunov Exponent** | 5 | 9 | 2 | 3 | 8 | 1 | 6 | 7 | 4 | 10 |

For measurement values, ApEn values are the same for cases sharing similar patterns and similar values. This is also seen in the SampEn and Maximal Lyanupov Exponent methods. For example, as shown in Figure 13, pilot Case 5 and Case 9 have similar weekly PPC reports, which means they have a similar performance pattern. Each individual algorithm reflects the same complexity level for similar construction performance.

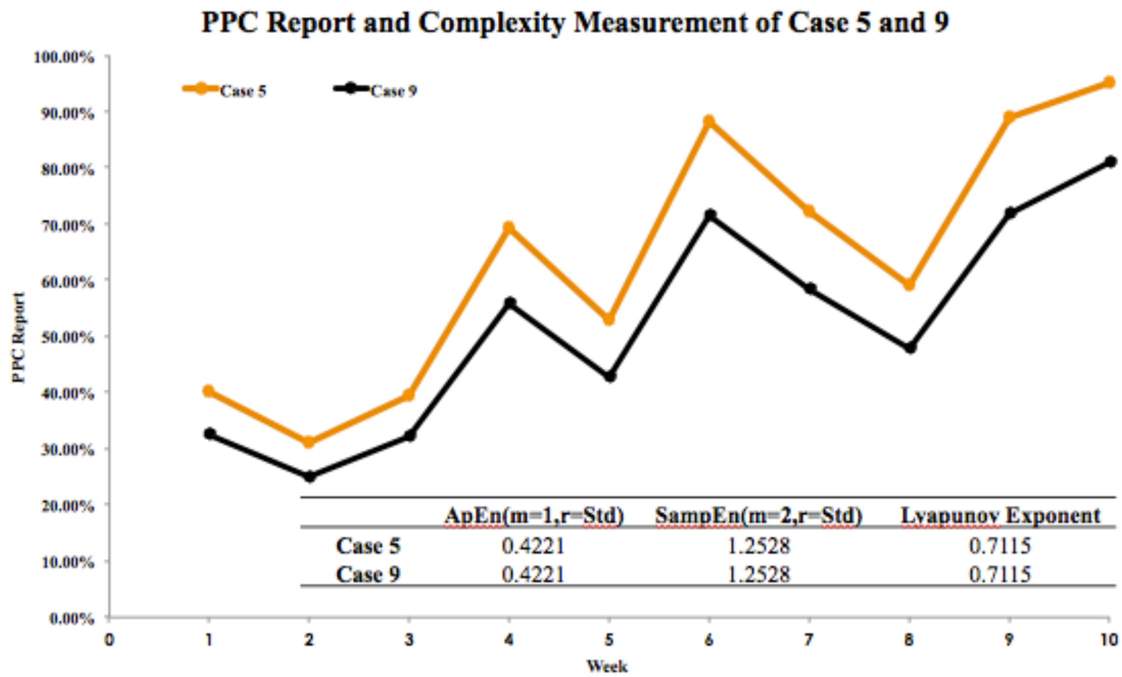



Figure 13 PPC Report and Complexity Measurement of Case 5 and 9

*5.5 Normalization of Complexity Measurements and Analysis*

For the different entropy algorithms based on different principles, or for the same algorithms with different parameters, it makes no sense to compare their measured results directly. Normalization would be a necessary next step before comparing results from different proposed methods. In this study, normalization of ratings or measurements means adjusting values measured on different scales or different algorithms to a common scale. In order to normalize different measurements from the proposed entropy algorithms, each group of complexity measurements will be divided by the maximal measurement in the group in order to get the common scale, which ranges from 0 to 1. The normalized results of the pilot case analysis are listed in Table 16 and plotted in Figure 14.

Table 16 Normalized Results of Complexity Measurement for Pilot Cases

| **Proposed Methods** | **CASE 1** | **CASE 2** | **CASE 3** | **CASE 4** | **CASE 5** | **CASE 6** | **CASE 7** | **CASE 8** | **CASE 9** | **CASE 10** |
|---|---|---|---|---|---|---|---|---|---|---|
| ApEn(r=0.1*Std) | 0.6192 | 0.7236 | 0.7236 | 0.3428 | 0.3428 | 1.0000 | 0.6192 | 0.0000 | 0.3428 | 0.0663 |
| ApEn(r=0.2*Std) | 0.7110 | 0.7110 | 1.0000 | 0.3665 | 0.4790 | 0.8749 | 0.8567 | 0.2652 | 0.4790 | 0.0000 |
| ApEn(r=0.5*Std) | 0.9613 | 0.5933 | 0.7631 | 0.4138 | 0.5896 | 1.0000 | 1.0000 | 0.8715 | 0.5896 | 0.5130 |
| ApEn(r=Std) | 0.6233 | 0.8022 | 0.7314 | 0.4010 | 1.0000 | 0.6528 | 0.6397 | 0.6966 | 1.0000 | 0.4010 |
| SampEn(r=Std) | 0.2847 | 0.4692 | 0.8770 | 0.2542 | 1.0000 | 0.2296 | 0.1781 | 0.3752 | 1.0000 | 0.2542 |
| PEn (n=2) | 0.9266 | 1.0000 | 0.9266 | 1.0000 | 1.0000 | 0.9266 | 1.0000 | 1.0000 | 1.0000 | 0.9266 |
| PEn (n=3) | 0.6245 | 0.8623 | 0.9000 | 0.9000 | 0.9000 | 0.7623 | 1.0000 | 0.7623 | 0.9000 | 0.8623 |
| PEn (n=4) | 0.8982 | 0.8982 | 1.0000 | 1.0000 | 1.0000 | 0.8982 | 1.0000 | 0.8982 | 1.0000 | 0.8982 |
| Maximal Lyapunov Exponent | 0.4872 | 0.9610 | 0.8374 | 0.1741 | 1.0000 | 0.2953 | 0.1896 | 0.4883 | 1.0000 | 0.1687 |

Note: All ApEn in the form, compared runs m = 1; compared runs (m) for Sample is equal to 2 in the form.

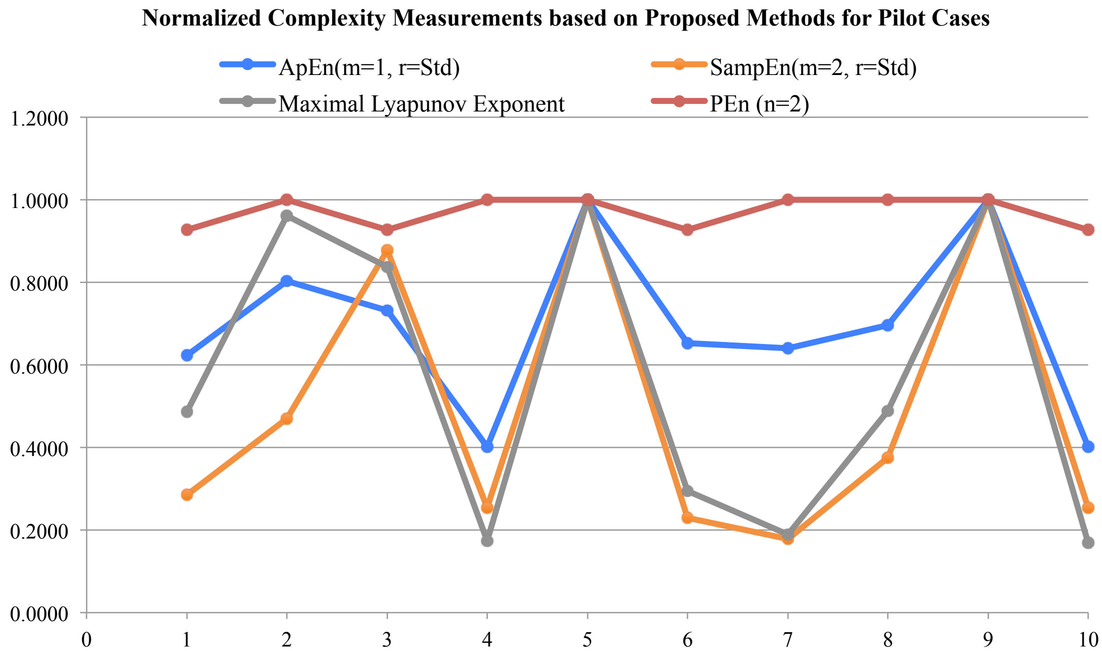



Figure 14 Normalized Complexity Measurements based on Proposed Methods for Pilot Cases

From Figure 14, we verify the conclusion made on the basis of the inferential analysis of the complexity measurements of each proposed method. Complexity measurements based on Permutation Entropy are not significant, and are mostly between 0.8 to 1; complexity ranks of the pilot cases based on Permutation Entropy also differ from the other proposed methods. The complexity ranks for pilot cases based on Approximate Entropy, Sample Entropy and the Maximal Lyapunov Exponent are similar to each other; they all rank Cases 5 and 9 as the most complex pilot cases. In addition, we see, for most of the cases, complexity measurements based on both Sample Entropy and the Maximal Lyapunov Exponent are pretty close.

## *5.6 Conclusion*

Based on the above analysis of results from the pilot cases study, the result and workability of each proposed entropy algorithms are summarized in Table 17. The results based on Six Sigma Analysis may be used as verification for the other proposed methods, because it only provides a rough ranking. Except for Permutation Entropy and Sample Entropy with narrow tolerance, which do not work for small sample size in close performance scenarios, the rest of the proposed algorithms are all workable for the pilot cases.

Table 17 Workability of Proposed Algorithms for Pilot Cases

| **Proposed Algorithms** | | **Workability** |
|---|---|---|
| **Description** | **Example** | (with Sample Size 10) |
| ApEn with Small Tolerance | ApEn(m=1, r=0.1*Std); ApEn(m=1, r=0.2*Std) | **Yes** |
| ApEn with Large Tolerance | ApEn(m=1, r=0.5*Std); ApEn(m=1, r=1*Std) | **Yes** |
| SampEn with Small Tolerance | SampEn(m=2, r=0.1*Std); SampEn(m=2, r=0.2*Std) | **No** |
| SampEn with Large Tolerance | SampEn(m=2, r=0. 5*Std); SampEn(m=2, r=1*Std) | **Yes** |
| PEn | PEn (n=2); PEn (n=3) | **No** |
| Maximal Lyapunov Exponent | | **Yes** |
| Six Sigma Analysis | | **Yes** |

Even though PPC is a comprehensive parameter to indicate the performance of a building construction project, it is still in its early stages. Among all construction projects and contractors, even though there are an increasing number adapting Lean Construction philosophy and methods, only a few have kept a complete PPC record for their work. The limited number of pilot cases are a direct result of this situation. However, with further simulation in the following chapters, these pilot cases and their simulated scenarios may still provide plenty of worthy conclusions for this advanced research of complexity measurement in building construction and for future studies.

# CHAPTER VI

# SIMULATION AND COMPARATIVE ANALYSIS

Pilot cases provide a limited number for complexity analysis in building construction systems. In order to test the proposed entropy algorithms in different scenarios and to make sure the proposed quantitative measuring model could be accepted in real construction work, a Monte Carlo simulation was used to generate more random numbers for each pilot case and to simulate different performance scenarios of the pilot cases. Based on these simulated cases, the complexity of construction could be measured and the entropy algorithms could be further compared.

At this stage, two rounds of simulation were conducted by MATLAB to generate more random numbers based on the pilot cases for the purpose of representing different circumstances in building construction performance so as to further compare the proposed entropy algorithms. The first round of simulations followed the same track of each pilot case to produce more random numbers for each entropy model based on sample size. Then, the second round of simulations provided each single pilot case different performance scenarios to test the performance of selected entropy algorithms and to determine how the measured results represented the performance change.

*6.1 Simulation Round I and Its Result Analysis*

The first round simulation generated more random numbers, following the same pattern of each pilot case, to test the workability of proposed entropy models in real complex long-term building construction systems. It also tested the workability of proposed entropy models with different sample sizes.

### 6.1.1 Simulation Method and Simulated Results

The first round of simulation generated additional random numbers based on the same pilot case pattern, because most of the entropy algorithms suggest using larger samples sizes to achieve stable performance. Nineteen (19) different probability distribution scenarios were tested for each pilot case in order to find the specific pattern with the highest fitness to generate random numbers.

Not only were common probability distributions (e.g., normal distribution) tested, other special probability distributions were tested in order to find the right pattern for each case with the highest fitness. Probability distributions, including Beta, Birnbaum-Saunders, Burr, Exponential, Extreme Value, Gamma, Generalized Extreme Value, Generalized Pareto, Inverse Gaussian, Log-Logistic, Logistic, Lognormal, Nakagami, Non-parametric, Normal, Rayleigh, Rician, t Location-Scale, and Weilbull, were tested for each pilot case. Taking Case 1 as an example, nineteen (19) probability distributions and their relative fitness are listed in Table 18. From the results, we see that Beta Distribution has the largest log-likelihood for pilot Case 1; and random numbers were generated for case 1 based on Beta Distribution.

Table 18 Probability Distribution and Likelihood for Case 1

| **Probability Distribution** | **Log Likelihood** | **Probability Distribution** | **Log Likelihood** |
|---|---|---|---|
| Beta | 0.9239 | Logistic | 0.1697 |
| Birnbaum-Saunders | -1.6250 | Lognormal | -1.6772 |
| Burr | 0 | Nakagaml | -0.0436 |
| Exponential | -4.1987 | Non-parametric | 0.1076 |
| Extreme Value | 0.3215 | Normal | 0.2476 |
| Gamma | -0.6968 | Rayleigh | -0.2406 |
| Generalized Extreme Value | 0.5802 | Rician | 0.3590 |
| Generalized Pareto | 0 | t Location-Scale | 0.2744 |
| Inverse Gaussian | -1.8040 | Weilbull | 0.1897 |
| Log-Logistic | -1.4851 | | |

According to the probability distribution pattern with the highest fitness, 50, 100 and 200 more numbers will be generated for each case for further comparative analysis of the proposed entropy algorithms. The first run (I-A) generated 50 more random numbers; each case then had 60 numbers total. The second run (I-B) generated 50 more random numbers than in Run I-A, making each case have 110 numbers total. The last run, which is Run I-C, generated 100 more random numbers than with Run I-B; each case then had 210 numbers total in the end. Each simulated case was tested by using the proposed entropy algorithms along with Six Sigma Analysis and the Maximal Lyapunov Exponent.

### 6.1.2 Descriptive Analysis of Round I Simulation and Comparison

In order to gain a better understanding of different complexity measurement models, the comparative analysis of the measured results and each algorithm were compared from two perspectives. Entropy algorithms were compared with each other in the same run and they were also compared for the same case across different runs. As most of the cases possessed similar probability distributions, the sigma level of the cases became ever closer while the sample size increased. As a result of this, for simulated cases in Run I, only proposed entropy algorithms and Maximal Lyapunov Exponent were compared and analyzed.

*Run I-A*

In Run I-A, each case had 60 weekly PPC values to represent its long-term performance. Proposed entropy algorithms with different parameter combinations were used to measure the complexity of its performance. Along with the results of Maximal Lyapunov Exponent, the measured results for Run I-A are listed in Table 19. For each individual entropy algorithm, the complexity level of each case increased with much stricter parameter selection. These stricter selections of parameters included low tolerance for ApEn and Sample, and more units in the compared template for PEn. As illustrated in Figure 15, when the tolerance widened (from 0.1 to 0.2 times the standard deviation), the cases became less complex.

Table 19 Complexity Measurements for Simulated Cases Round I Run A (I-A)

| Case | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|
| | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **1** | 1.3544 | 1.1998 | Inf | 1.9810 | 1.3782 | 0.7229 | 0.6861 | 1.7400 | 1.5519 |
| **2** | 1.4923 | 1.1835 | 1.3863 | 2.0794 | 1.4553 | 0.6867 | 0.6861 | 1.7016 | 1.6206 |
| **3** | 1.4742 | 1.1436 | Inf | Inf | 1.2448 | 0.6162 | 0.6815 | 1.7512 | 1.4418 |
| **4** | 1.4581 | 1.2354 | Inf | 1.9459 | 1.3365 | 0.5050 | 0.6919 | 1.7789 | 1.5576 |
| **5** | 1.5292 | 1.2652 | Inf | 1.8718 | 1.2104 | 0.7895 | 0.6919 | 1.7855 | 1.4652 |
| **6** | 1.3378 | 1.1925 | Inf | 2.5123 | 1.1680 | 0.5863 | 0.6896 | 1.7501 | 1.3570 |
| **7** | 1.5191 | 1.1497 | Inf | 2.0794 | 1.3618 | 0.6088 | 0.6930 | 1.7646 | 1.4690 |
| **8** | 1.6015 | 1.2995 | Inf | 1.7918 | 1.3257 | 0.6113 | 0.6919 | 1.7729 | 1.6958 |
| **9** | 1.5350 | 1.2205 | Inf | 2.8904 | 1.1827 | 0.7223 | 0.6919 | 1.7533 | 1.7302 |
| **10** | 1.4713 | 1.2195 | Inf | 2.4423 | 1.0684 | 0.4796 | 0.6896 | 1.7759 | 1.5394 |

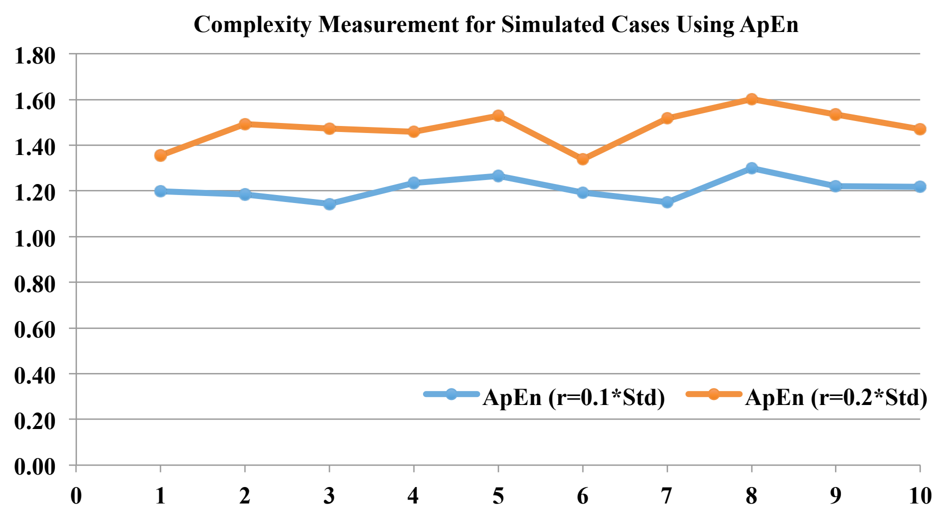



Figure 15 Complexity Measurement for Simulated Cases Using ApEn

From the results listed in Table 19, we concluded that the narrower tolerance of Sample Entropy would not work well with a sample size of approximately 50. Plotting all the measurement results visually (see Figure 16), we see that the complexity of the ten cases were not significantly different based on each method, especially for Permutation Entropy. Because each method has its own perspective, similar cases were ranked differently. However, the ranks of the cases based on Sample Entropy is still similar to the rank based on Maximal Lyapunov Exponent, as observed in the pilot case analysis.

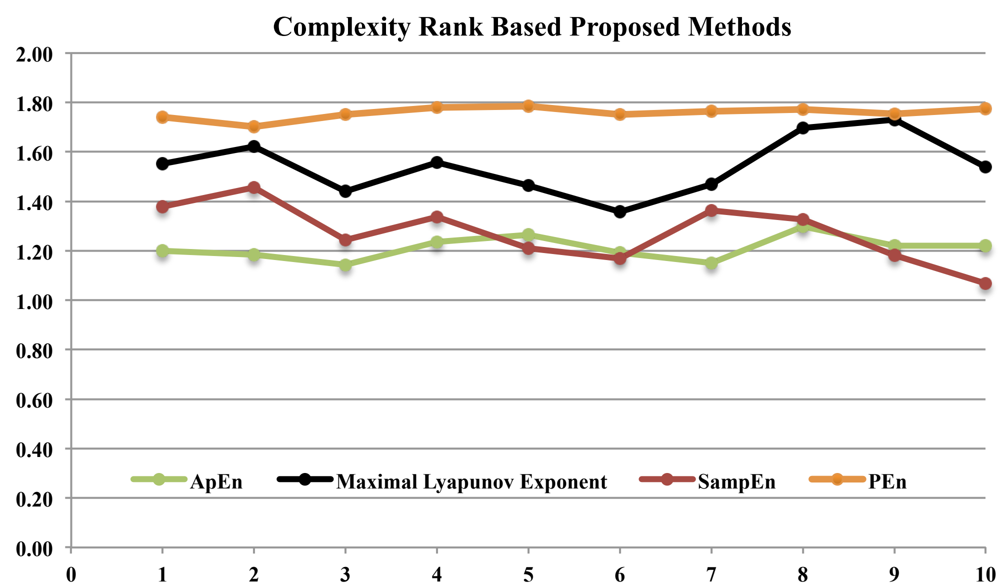



Figure 16 Complexity Rank Based Proposed Methods

*Run I-B*

In Run I-B, more random numbers were generated, based on Run I-A, with 110 PPC values for each case to represent the long-term performance of building construction. In addition, the increased sample size was used to test the calculation speed of each proposed entropy model. Like those conducted in the first run, the proposed entropy algorithms with the same parameter combinations and Maximal Lyapunov Exponent were tested in Run I-B again. Their results are shown in Table 20. From the horizontal comparison of different methods in the single run, we see that stricter parameter selections resulted in higher complexity levels for the same case using the same entropy algorithm. For ApEn, in Figure 17, we see that the complexity measurement increased when we chose smaller tolerance. For SampEn in Figure 18 (more than 100 numbers), a narrower tolerance provided precise complexity measurements, while wider tolerances became less sensitive to similar data patterns. This lower sensitivity was also found in the measurement of PEn.

Table 20 Complexity Measurement for Simulated Cases Round I Run B (I-B)

| Case | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|
| | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **1** | 1.7446 | 1.6203 | 3.2958 | 2.0541 | 1.3083 | 0.6939 | 0.6880 | 1.7658 | 1.8105 |
| **2** | 1.7828 | 1.6626 | 1.5581 | 1.9299 | 1.2144 | 0.6817 | 0.6880 | 1.7270 | 1.6504 |
| **3** | 1.7710 | 1.6593 | Inf | 3.4812 | 1.3194 | 0.6573 | 0.6911 | 1.7824 | 1.7661 |
| **4** | 1.7256 | 1.642 | 2.6391 | 1.8971 | 1.3031 | 0.5715 | 0.6921 | 1.7788 | 1.686 |
| **5** | 1.6644 | 1.6288 | 2.6391 | 2.5200 | 1.3297 | 0.7246 | 0.6911 | 1.7786 | 1.7015 |
| **6** | 1.6593 | 1.6141 | 3.2189 | 1.9924 | 1.1109 | 0.5744 | 0.6921 | 1.7658 | 1.6547 |
| **7** | 1.7393 | 1.6607 | Inf | 2.2900 | 1.3427 | 0.6571 | 0.6921 | 1.7845 | 1.9201 |
| **8** | 1.7086 | 1.5974 | 3.0910 | 1.9889 | 1.3463 | 0.6104 | 0.6931 | 1.7768 | 1.6707 |
| **9** | 1.8049 | 1.7151 | 2.4849 | 2.5337 | 1.2605 | 0.7128 | 0.6928 | 1.7866 | 1.8773 |
| **10** | 1.7330 | 1.6997 | 3.0445 | 1.7819 | 1.0420 | 0.5427 | 0.6931 | 1.7821 | 1.6914 |

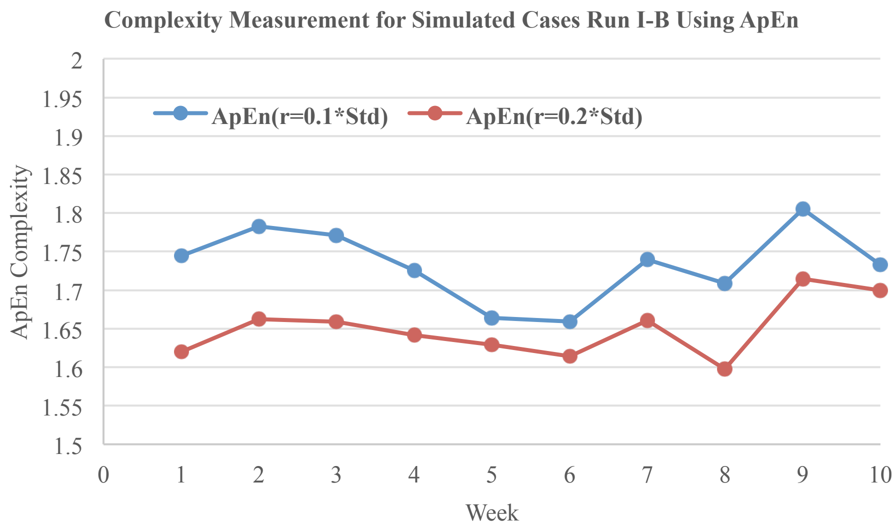



**Figure 17 Complexity Measurement for Simulated Cases Run I-B Using ApEn**

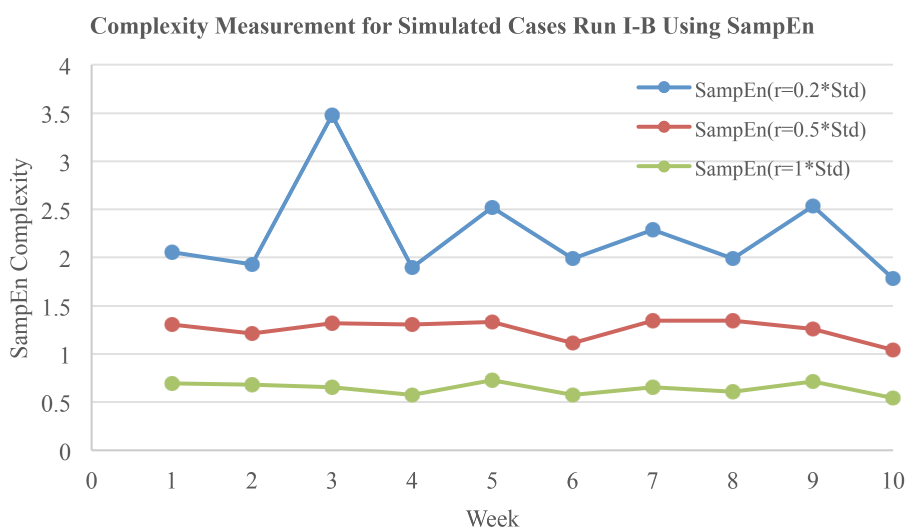



Figure 18 Complexity Measurement for Simulated Cases Run I-B Using SampEn

For further analysis, we plotted the measurement results with the stricter criteria, along with the results of Maximal Lyapunov Exponent (see Figure 19) for an intuitive comparison. As shown in the chart, Sample Entropy with a narrower tolerance was sensitive to the similar data patterns of larger sample sizes, compared to Approximate Entropy and Permutation Entropy. The rank fluctuation based on Sample Entropy is also similar to the rank based on Maximal Lyapunov Exponent as plotted in Figure 19.

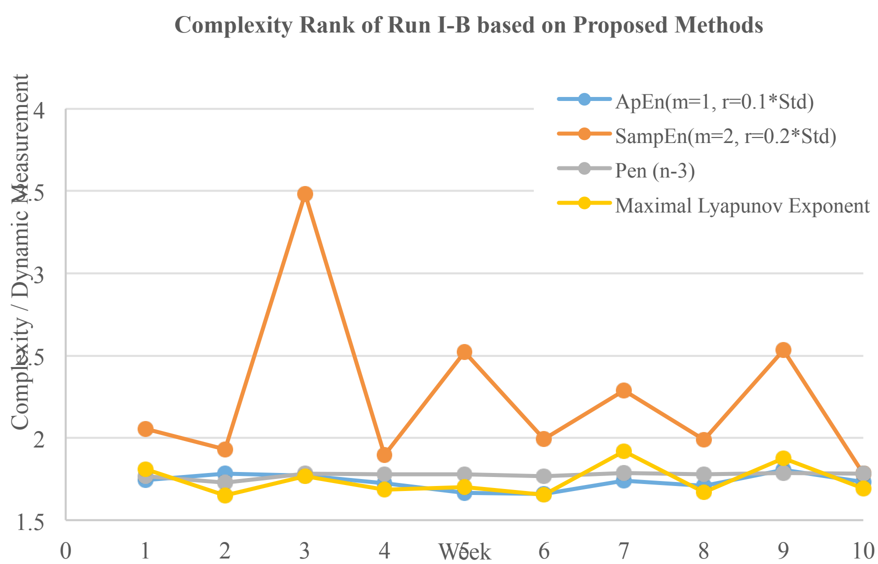



Figure 19 Complexity Rank of Run I-B based on Proposed Methods

*Run I-C*

In the last run of round I simulation, Run I-C, 100 more random numbers were generated by using Monte Carlo simulation based on the cases in the previous run. At this point, each case had 210 PPC values for complexity measurements. The proposed entropy algorithms with different parameter combinations and Maximal Lyapunov Exponents were tested in Run I-C again. The results are shown in Table 21. Based on the horizontal comparison of different methods in the last run of the first round simulation, for each specific entropy model, stricter parameter selections resulted in higher complexity levels for the same case.

For ApEn, in Figure 20, we see that the complexity measurement increased for the same case when we chose smaller tolerances. However, with the large sample size, actually most of the simulated cases followed a similar pattern. As a result of this, we see that the ApEn complexity measurement for these cases is more or less the same. For example, the complexity measurement based on Approximate Entropy Model, with m=1 and r= 0.1 times the standard deviation of the sample, falls within the range of (1.95, 2.1). For SampEn in Figure 21, with a larger sample size, it actually worked better than with small sample sizes, like in the pilot cases and in the first run of Round I simulation. With a narrower tolerance, it provided precise complexity measurements, while wider tolerances became less sensitive to similar data patterns just like with ApEn. This character of less sensitivity could still be found in the measurement of PEn.

Table 21 Complexity Measurement for Simulated Cases Round I Run C (I-C)

| Case | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|
| | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **1** | 2.0084 | 1.9585 | 3.0204 | 2.3820 | 1.3298 | 0.6823 | 0.6906 | 1.7758 | 1.9721 |
| **2** | 1.9710 | 1.9003 | 2.2900 | 1.9763 | 1.3379 | 0.6847 | 0.6890 | 1.7657 | 2.0150 |
| **3** | 1.9929 | 1.8861 | 2.9444 | 2.0422 | 1.2550 | 0.6200 | 0.6912 | 1.7842 | 2.1142 |
| **4** | 1.9952 | 1.9250 | 2.7344 | 2.0857 | 1.2483 | 0.6269 | 0.6922 | 1.7860 | 1.8944 |
| **5** | 2.0690 | 1.9480 | 3.4177 | 2.2336 | 1.3074 | 0.6794 | 0.6930 | 1.7904 | 2.0274 |
| **6** | 2.0097 | 1.9064 | 2.4532 | 2.1053 | 1.2274 | 0.6337 | 0.6906 | 1.7754 | 2.0757 |
| **7** | 2.0945 | 1.9959 | 2.7568 | 2.2274 | 1.2876 | 0.6585 | 0.6930 | 1.7911 | 2.1029 |
| **8** | 2.0252 | 1.8897 | 2.8332 | 1.9027 | 1.1805 | 0.5870 | 0.6930 | 1.7849 | 1.9440 |
| **9** | 2.0322 | 1.9213 | 2.5953 | 2.0541 | 1.3184 | 0.6969 | 0.6918 | 1.7787 | 2.0932 |
| **10** | 2.0346 | 1.9232 | 2.6288 | 1.8142 | 1.0276 | 0.5423 | 0.6929 | 1.7843 | 1.9983 |

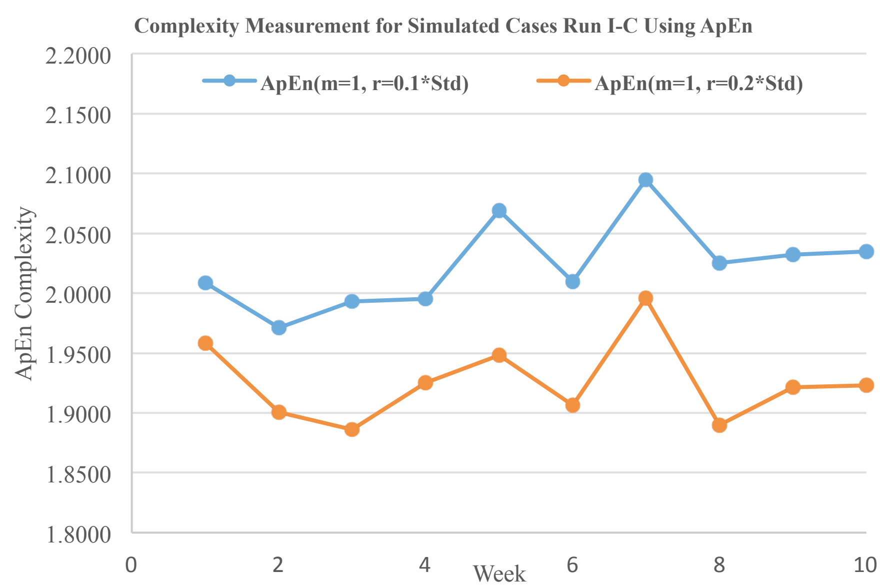



Figure 20 Complexity Measurement for Simulated Cases Run I-C Using ApEn

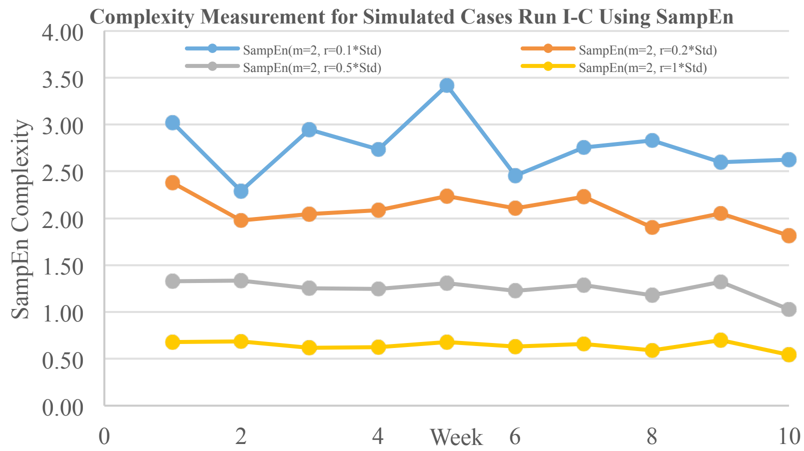



Figure 21 Complexity Measurement for Simulated Cases Run I-C Using SampEn

In the end, we plotted the measurement results with the stricter criteria, which were ApEn with m=1 and r=0.1*Std, SampEn with m=2 and r=0.1 Std, and PEn with n=3, along with the results of Maximal Lyapunov Exponent. See Figure 22 for an intuitive comparison. As shown in the chart, Sample Entropy with a narrower tolerance is sensitive to the similar data pattern with larger sample sizes compared with Approximate Entropy, Permutation Entropy and Maximal Lyapunov Exponent.

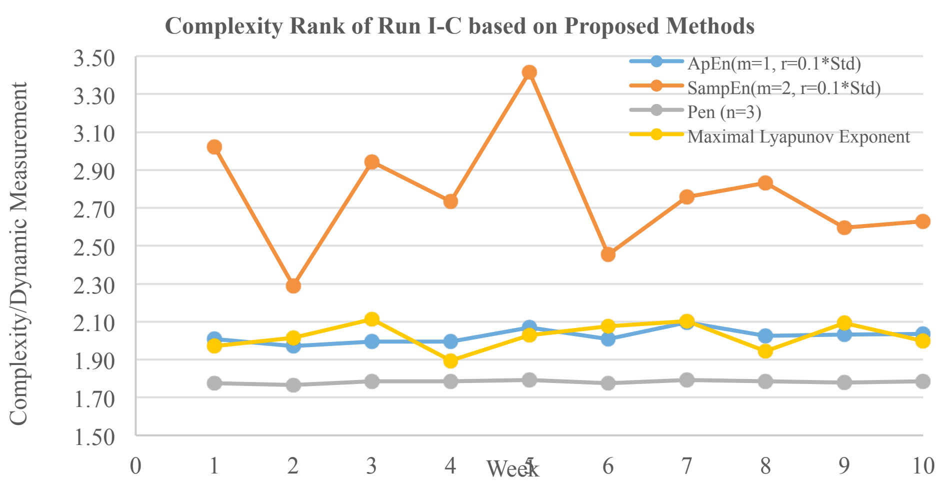



Figure 22 Complexity Rank of Run I-C based on Proposed Methods

*Longitudinal Comparison of Complexity Measurements for Single Case in Three Different Runs*

From the horizontal comparison of complexity measurements of the ten simulated cases in the same run, we note that with the stricter parameter selection, proposed entropy algorithms became more sensitive, even for the cases with similar probability distribution patterns. However, with the larger sample size, most of the pilot cases started to show a similar pattern in the complexity measurement results. Most of the proposed entropy algorithms began to fall within a small range, while only the Sample Entropy with narrow tolerance still provided precise complexity measurements to tell the difference among those cases. However, Approximate Entropy or Sample Entropy with wider tolerances worked better for the smaller sample size.

For further analysis and comparison of the proposed entropy algorithms, they were also compared longitudinally for the same case with a different sample size. In order to conduct this further comparison, we randomly chose Case 2 as an example to illustrate the analysis procedure and results; for other cases, the analysis procedure and results are the same.

The complexity measurement results for Case 2 in three different runs are presented in Table 22. From the results, one can conclude that in the runs with stricter parameter selection criteria (narrower tolerance or more units in the compared template), higher complexity measurements are obtained. For the Entropy Algorithms, especially the Sample Entropy, the measurement results became more sensitive with the stricter parameters.

Table 22 Complexity Measurements for Case 2 in Different Simulated Runs

| Case | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|
| | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **Case 2-A** | 1.4923 | 1.1835 | 1.3863 | 2.0794 | 1.4553 | 0.6867 | 0.6861 | 1.7016 | 1.6206 |
| **Case 2-B** | 1.7828 | 1.6626 | 1.5581 | 1.9299 | 1.2144 | 0.6817 | 0.688 | 1.727 | 1.6504 |
| **Case 2-C** | 1.9710 | 1.9003 | 2.2900 | 1.9763 | 1.3379 | 0.6847 | 0.6890 | 1.7657 | 2.0150 |

From the data in Table 22, we also checked the complexity track of each case with different simulated sample sizes. For Case 2, the three different simulated PPC values are plotted in Figure 23. The latter run of each case, based on its previous run to generate more random numbers, follows the same probability distribution pattern.

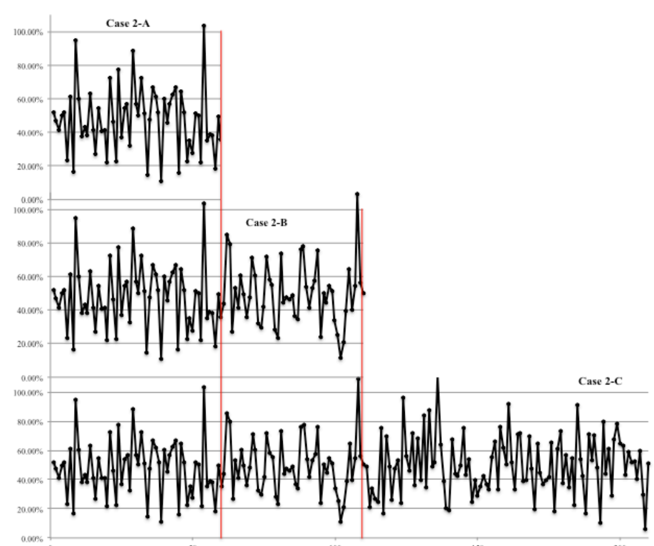



Figure 23 Weekly PPC Record for Case 2 in Three Simulated Run

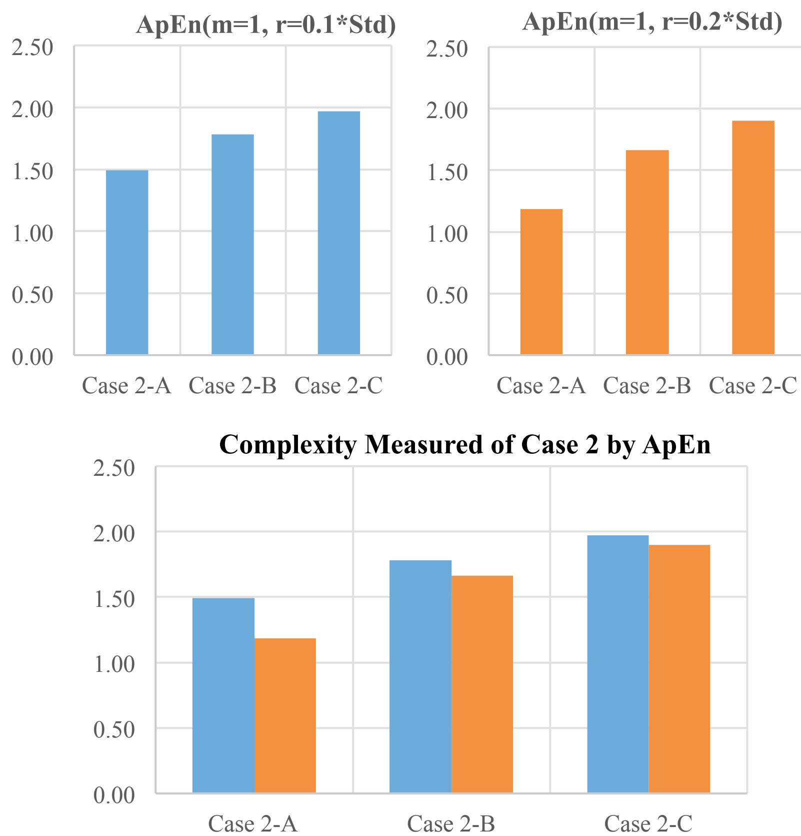



Figure 24 Complexity Measurements of Simulated Case 2 Using ApEn

From the complexity measurement for Case 2 using ApEn, as shown in Figure 24, the complexity level of the cases increased by using both tolerance choices, which means that with more random numbers being generated, the case became more complex. Also, the SampEn complexity measurement results, as presented in Figure 25, show that the same complexity increase was observed when the tolerance was equal to 0.1 times the standard deviation of the simulated case, which is the smallest tolerance. However, for the other tolerances, the wider the tolerance, the less sensitive it is to the complexity level. For the largest tolerance, which is equal to the standard deviation of the simulated sample, the complexity measurement was almost the same for the three runs. From the measurement results for PEn, shown in Figure 26, the lower sensitivity was also found.

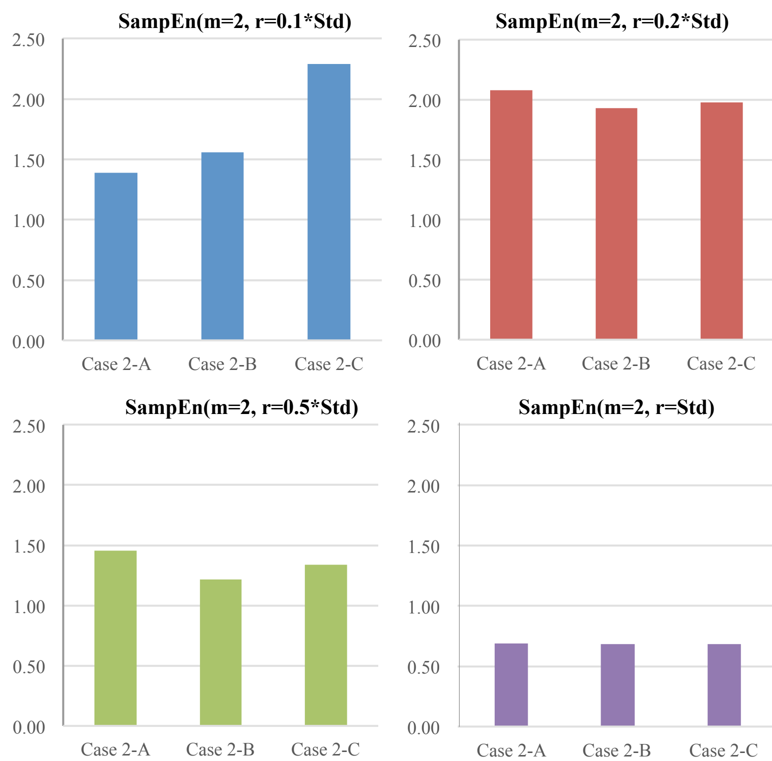



Figure 25 Complexity Measurements of Simulated Case 2 Using SampEn

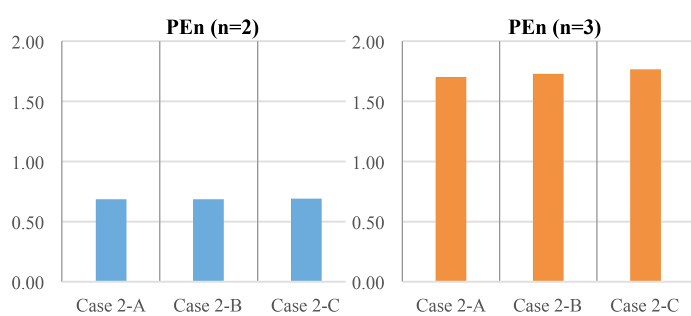



Figure 26 Complexity Measurements of Simulated Case 2 Using PEn

In the end, with the Maximal Lyapunov Exponent method (see Figure 27), the case's complexity level also increased with the generation of random numbers.

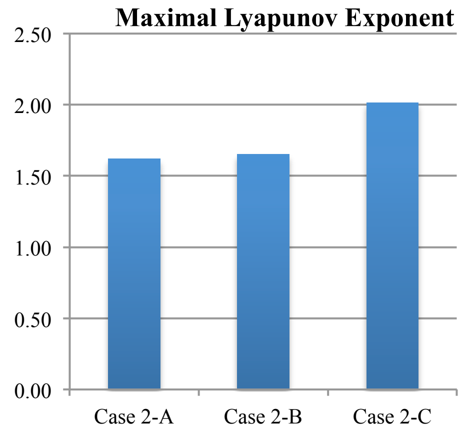



Figure 27 Complexity of Simulated Case 2 ranked by Maximal Lyapunov Exponent

As discussed above, all the observed rankings and results of proposed entropy algorithms for simulated Case 2 are also seen in the results of the other simulated cases. With the proof of the longitudinal comparison of all the simulated cases in Round I, the results indicate that with narrower tolerances, both Approximate Entropy and Sample Entropy show increasing complexity with the more and more random numbers of each pilot case. This increasing trend is also observed when the Maximal Lyapunov Exponent is applied. Permutation Entropy and Sample Entropy with larger tolerances are not very sensitive with the sample number of the simulated cases based on the same pilot case.

### 6.1.3 Normalization of Complexity Measurements in Round I Simulation

As mentioned in the pilot cases analysis, directly comparing the complexity measurements based on different entropy algorithms is not an appropriate perspective. To overcome this problem, normalization was conducted in order to compare different entropy complexity results. The method we used to normalize the measurements is the one we used and tested in the pilot cases analysis, which is to divide the maximal measurement from each group of results.

The normalized complexity measurements for simulation Run I-A are listed in Table 23, and plotted in Figure 28. Permutation Entropy results are the least significant, while the results based on Sample Entropy are the most significant. Except for Case 9, the complexity trend based on Sample Entropy and the Maximal Lyapunov Exponent are similar. All of these verified the conclusions we found in the previous analysis of Run I-A.

Table 23 Normalized Complexity Measurements for Simulation Run I-A

| Case | ApEn (m=1) | | SampEn (m=2) | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|
| | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **1** | 0.8457 | 0.9233 | 0.9470 | 0.9156 | 0.9900 | 0.9745 | 0.8969 |
| **2** | 0.9318 | 0.9107 | 1.0000 | 0.8698 | 0.9900 | 0.9530 | 0.9367 |
| **3** | 0.9205 | 0.8800 | 0.8554 | 0.7805 | 0.9834 | 0.9808 | 0.8333 |
| **4** | 0.9105 | 0.9507 | 0.9184 | 0.6396 | 0.9984 | 0.9963 | 0.9002 |
| **5** | 0.9549 | 0.9736 | 0.8317 | 1.0000 | 0.9984 | 1.0000 | 0.9002 |
| **6** | 0.8353 | 0.9177 | 0.8026 | 0.7426 | 0.9951 | 0.9802 | 0.7843 |
| **7** | 0.9485 | 0.8847 | 0.9358 | 0.7711 | 1.0000 | 0.9883 | 0.8490 |
| **8** | 1.0000 | 1.0000 | 0.9109 | 0.7743 | 0.9984 | 0.9929 | 0.9223 |
| **9** | 0.9585 | 0.9392 | 0.8127 | 0.9149 | 0.9984 | 0.9820 | 1.0000 |
| **10** | 0.9187 | 0.9384 | 0.7341 | 0.6075 | 0.9951 | 0.9946 | 0.8897 |

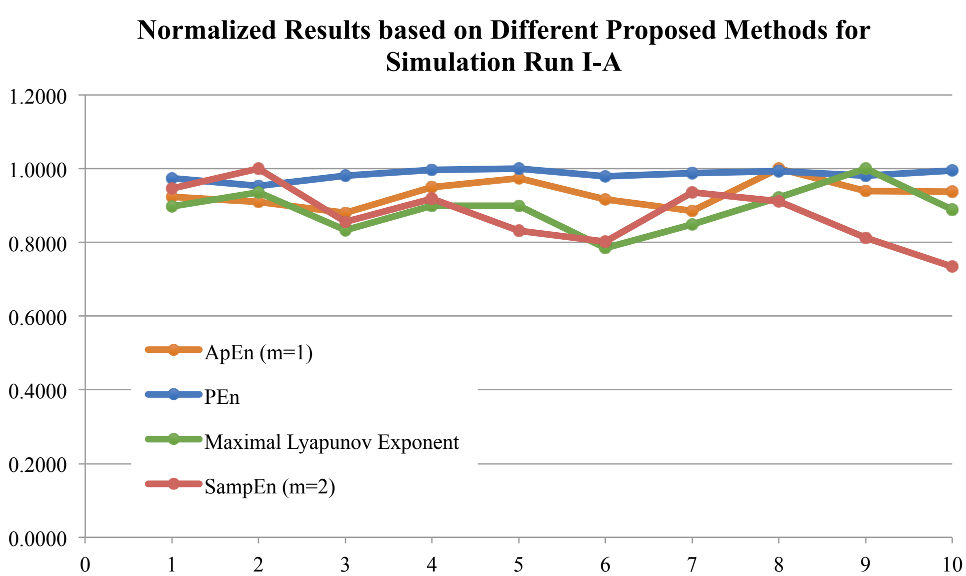



Figure 28 Normalized Results based on Different Proposed Methods for Simulation I-A

The normalized complexity measurements for simulation Run I-B are listed in Table 24, and plotted in Figure 29. The normalized results indicate that Sample Entropy with a narrower tolerance is the most sensitive method for a similar data pattern with a larger sample size, compared with Approximate Entropy, Permutation Entropy and even the

Maximal Lyapunov Exponent. The rank fluctuation based on Sample Entropy is also similar to the rank based on Maximal Lyapunov Exponent. All of these statements prove the previous conclusions based on the descriptive analysis of Run I-B results.

Table 24 Normalized Complexity Measurements for Simulation Run I-B

| Case | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|
| | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **1** | 0.9666 | 0.9447 | N/A | 0.5901 | 0.9718 | 0.9576 | 0.9926 | 0.9884 | 0.9429 |
| **2** | 0.9878 | 0.9694 | N/A | 0.5544 | 0.9020 | 0.9408 | 0.9926 | 0.9666 | 0.8595 |
| **3** | 0.9812 | 0.9675 | N/A | 1.0000 | 0.9800 | 0.9071 | 0.9971 | 0.9976 | 0.9198 |
| **4** | 0.9561 | 0.9574 | N/A | 0.5450 | 0.9679 | 0.7887 | 0.9986 | 0.9956 | 0.8781 |
| **5** | 0.9222 | 0.9497 | N/A | 0.7239 | 0.9877 | 1.0000 | 0.9971 | 0.9955 | 0.8862 |
| **6** | 0.9193 | 0.9411 | N/A | 0.5723 | 0.8252 | 0.7927 | 0.9986 | 0.9884 | 0.8618 |
| **7** | 0.9637 | 0.9683 | N/A | 0.6578 | 0.9973 | 0.9068 | 0.9986 | 0.9988 | 1.0000 |
| **8** | 0.9466 | 0.9314 | N/A | 0.5713 | 1.0000 | 0.8424 | 1.0000 | 0.9945 | 0.8701 |
| **9** | 1.0000 | 1.0000 | N/A | 0.7278 | 0.9363 | 0.9837 | 0.9996 | 1.0000 | 0.9777 |
| **10** | 0.9602 | 0.9910 | N/A | 0.5119 | 0.7740 | 0.7490 | 1.0000 | 0.9975 | 0.8809 |

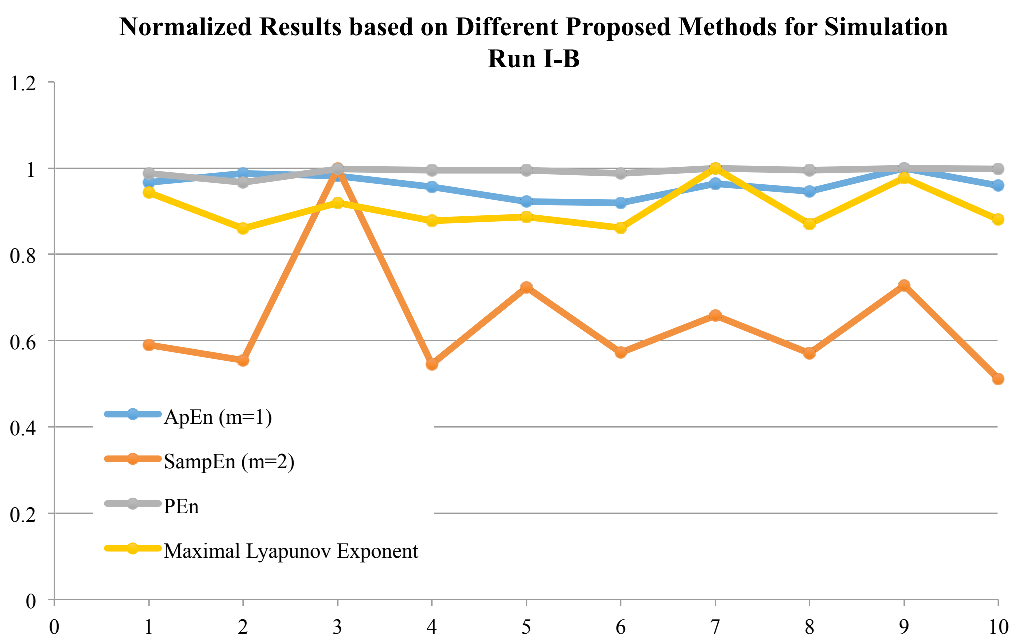



Figure 29 Normalized Results based on Different Proposed Methods for Simulation I-B

The normalized complexity measurements for simulation Run I-C are listed in Table 25, and plotted in Figure 30. As in Figure 22 in the previous analysis of simulation Run I-C, the normalized complexity measurements of simulation Run I-C also shows that the Sample Entropy with narrower tolerance offers the most significant results for a similar data pattern with larger sample size, as compared with Approximate Entropy, Permutation Entropy and Maximal Lyapunov Exponent.

Table 25 Normalized Complexity Measurements for Simulation Run I-C

| Case | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|
| | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **1** | 0.9589 | 0.9813 | 0.8838 | 1.0000 | 0.9939 | 0.9791 | 0.9965 | 0.9915 | 0.9328 |
| **2** | 0.9410 | 0.9521 | 0.6700 | 0.8297 | 1.0000 | 0.9825 | 0.9942 | 0.9858 | 0.9531 |
| **3** | 0.9515 | 0.9450 | 0.8615 | 0.8573 | 0.9380 | 0.8897 | 0.9974 | 0.9961 | 1.0000 |
| **4** | 0.9526 | 0.9645 | 0.8001 | 0.8756 | 0.9330 | 0.8996 | 0.9988 | 0.9972 | 0.8960 |
| **5** | 0.9878 | 0.9760 | 1.0000 | 0.9377 | 0.9772 | 0.9749 | 1.0000 | 0.9996 | 0.9589 |
| **6** | 0.9595 | 0.9552 | 0.7178 | 0.8838 | 0.9174 | 0.9093 | 0.9965 | 0.9912 | 0.9818 |
| **7** | 1.0000 | 1.0000 | 0.8066 | 0.9351 | 0.9624 | 0.9449 | 1.0000 | 1.0000 | 0.9947 |
| **8** | 0.9669 | 0.9468 | 0.8290 | 0.7988 | 0.8824 | 0.8423 | 1.0000 | 0.9965 | 0.9195 |
| **9** | 0.9703 | 0.9626 | 0.7594 | 0.8623 | 0.9854 | 1.0000 | 0.9983 | 0.9931 | 0.9901 |
| **10** | 0.9714 | 0.9636 | 0.7692 | 0.7616 | 0.7681 | 0.7782 | 0.9999 | 0.9962 | 0.9452 |

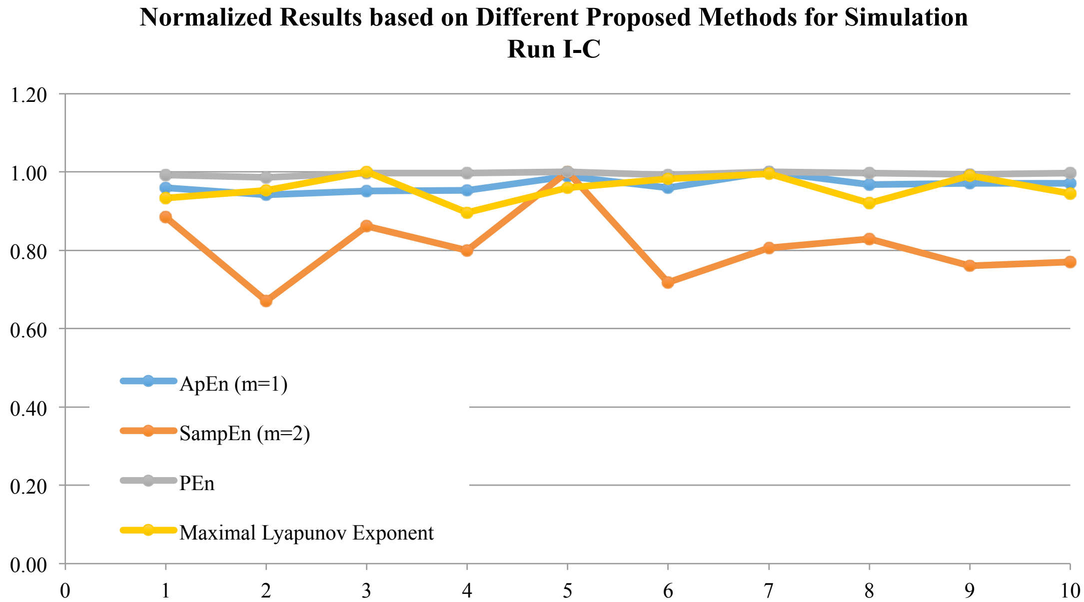



Figure 30 Normalized Results based on Different Proposed Methods for Simulation I-C

### 6.1.4 Inferential Analysis of Comparative Analysis for Round I

The Round I simulation with three different runs tested the workability of proposed entropy algorithms for different sample sizes based on similar performance patterns. The conclusions are based on analyses from different perspectives, including comparative analysis of ten (10) different cases in the same simulated run, and the comparison of the three different runs for the same case. All these conclusions have been verified by using normalized complexity measurements based on different proposed methods, as well. The workability of dealing with different sample sizes for each method is summarized in Table 26. From the summary, we conclude that Approximate Entropy and the Maximal Lyapunov Exponent only work well for sample sizes from 10 to 100. Only Sample

Entropy works for both small and large sample sizes, with the only change being the choice of appropriate tolerance. The larger the sample size , the narrower the tolerance should be for Sample Entropy. One other interesting outcome is that the rank based on Sample Entropy is similar to the rank based on the Maximal Lyapunov Exponent; both of these two methods could be used to obtain the complexity measurements.

Table 26 Proposed Methods' Workability of Dealing with Different Sample Sizes

| **Proposed Algorithms** | **Workability for Different Sample Size** | | | |
|---|---|---|---|---|
| | **60** | **110** | **210** | **Increasing Size** |
| ApEn with Small Tolerance | Yes | No | No | Yes |
| SampEn with Small Tolerance | No | Yes | Yes | Yes |
| SampEn with Large Tolerance | Yes | No | No | No |
| PEn | No | No | No | No |
| Maximal Lyapunov Exponent | Yes | Yes | No | Yes |
| Six Sigma Analysis | No | No | No | No |

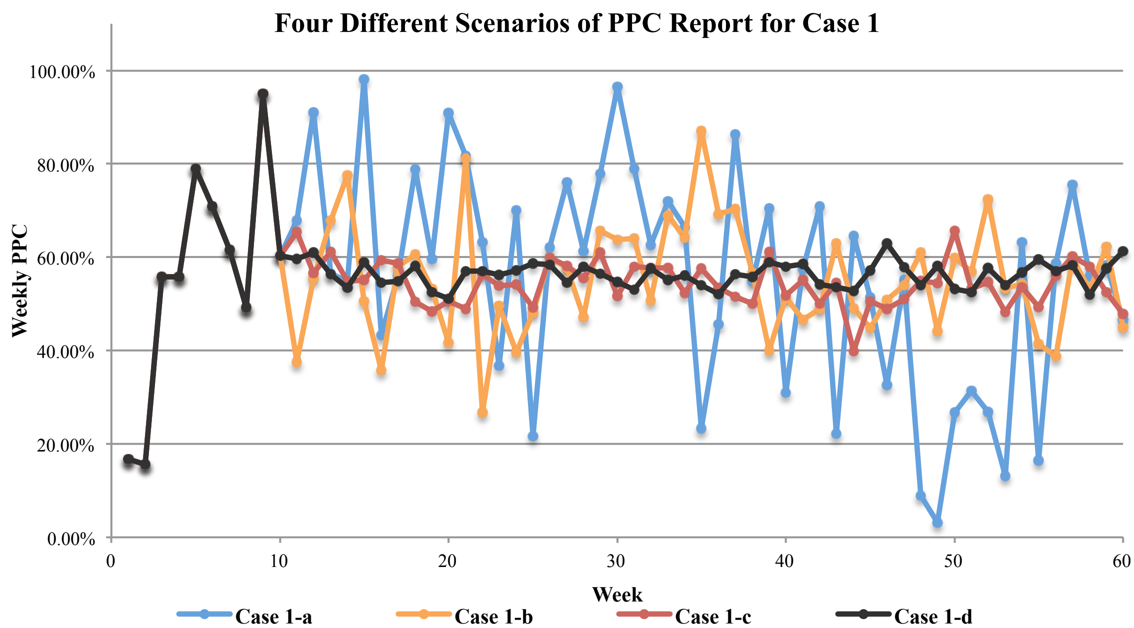
### *6.2 Simulation Round II and Its Results Analysis*

The first round simulation tested the workability and sensitivity of each proposed entropy algorithm based on different sample sizes of similar performance that resulted from three runs of Monte Carlo simulation for each pilot case. For each case, all random numbers were generated based on the same performance pattern, which was decided by the PPC probability distribution in the process of simulation. The second round of simulation generated more performance patterns to further test the proposed entropy algorithms. In round two simulations, the three runs (II-A, II-B, and II-C) were conducted: with additional random numbers (50, 100, and 200, respectively). However, in each run, the measurement results were not compared across different cases, instead, they were compared based on different performance patterns for the same case. For each pilot case, four different simulated scenarios were generated. Simulated random numbers fell in the range of (Mean-3*Standard Deviation, Mean-3*Standard Deviation) for scenario a of each case, (Mean-1.5*Standard Deviation, Mean-1.5*Standard Deviation) for scenario b, (Mean-0.5*Standard Deviation, Mean-0.5*Standard Deviation) for scenario c and (Mean-0.25*Standard Deviation, Mean-0.25*Standard Deviation) for the last scenario, which is scenario d. This time, the proposed entropy algorithms, along with Six Sigma Analysis and the Maximal Lyapunov Exponent, were tested on each scenario.

### 6.2.1 Simulation Method and Simulated Results

In round two simulation, three run will be conducted like in previous simulation that Run II-A with 50 more number, Run II-B with 100 more number and Run II-C with 200 number. However, in each run, the measurement results with not be compared across different cases, they will be compared based on different performance patterns for the same case. And for each pilot case, four different simulated scenarios will be generated. Simulated random numbers will fall in the range of (Mean-3*Standard Deviation, Mean-3*Standard Deviation) for scenario a of each case, (Mean-1.5*Standard Deviation, Mean-1.5*Standard Deviation) for scenario b, (Mean-0.5*Standard Deviation, Mean-0.5*Standard Deviation) for scenario c and (Mean-0.25*Standard Deviation, Mean-

0.25*Standard Deviation) for the last scenario, which is scenario d. For scenario a, if the lower boundary of the range of (Mean-3*Standard Deviation, Mean-3*Standard Deviation) was less than 0%, it was replaced by 0%; if the upper boundary was larger than 100%, if was replaced by 100% to make sure the PPC report would be no greater than 100% or less than 0%. Taking Pilot Case 1 as an example, the four different scenarios generated in Run A of simulation Round II are shown in Figure 31.




| Case 1-a | Case 1-b | Case 1-c | Case 1-d |
|---|---|---|---|
| (0,100%) | (19%,93%) | (44%,68%) | (50%,62%) |

Figure 31 Four Different Scenarios of PPC Report for Case 1

This time, the proposed entropy algorithms along with Six Sigma Analysis and the Maximal Lyapunov Exponent were tested for each scenario. The parameter selections for ApEn, SampEn and PEn will be the same with parameters used in simulation Round I. They are ApEn (m=1, r=0.1*Std), ApEn (m=1, r=0.2*Std), SampEn (m=2, r=0.1*Std), SampEn (m=2, r=0.2*Std), SampEn (m=2, r=0.5*Std), SampEn (m=2, r=1*Std), PEn (n=2) and PEn (n=3).

### 6.2.2 Descriptive Analysis of Round II Simulation and Comparison

In order to have a better understanding of all three proposed entropy algorithms, the measured results and each algorithm were compared from two perspectives. The entropy algorithms were compared with each other in the same run for the same case with different scenarios, and they were also compared for the same case and same scenario across different runs. From the comparative results, the ability to determine different construction performance levels were tested for each entropy algorithm, as was their ability to deal with different sample sizes.

*Run II-A*

In Run A of the second round simulation, each simulated pilot case contained 60 weekly PPC values to represent its long-term performance. In addition, compared with Run A in simulation Round I, each case had four different performance scenarios to simulate different project circumstances.

Six Sigma Analysis was first applied to each simulated case to rank the four simulated performances. The criteria of Six Sigma Analysis was the same as the criteria discussed in Chapter 5.1, which was the mean of each pilot case minus its standard deviation and that mean plus its standard deviation. The sigma levels of four scenarios for each pilot case are shown in Table 27 as the tested ranks for the measurement results of the proposed entropy algorithms.

The proposed entropy algorithms with different parameter combinations were used to measure the complexity of each scenario, and then compared with the other scenarios for the same case. Along with the results of the Maximal Lyapunov Exponent, the measured results for Run II-A are listed in Table 28.

Table 27 Six Sigma Analysis for Simulated Cases Round II Run A (II-A)

| Case | Scenario a | Scenario b | Scenario c | Scenario d |
|---|---|---|---|---|
| Yield | 68.33% | 90.00% | 95.00% | 95.00% |
| Sigma Level | 3 | 3 | 4 | 4 |

Table 28 Complexity Measurement for Simulated Cases Round II Run A (II-A)

| **Case** | | **ApEn (m=1)** | | **SampEn (m=2)** | | | | **PEn** | | **Maximal Lyapunov Exponent** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **r=0.1*Std** | **r=0.2*Std** | **r=0.1*Std** | **r=0.2*Std** | **r=0.5*Std** | **r=Std** | **n=2** | **n=3** | |
| **1** | a | 1.4957 | 1.2465 | Inf | 2.7726 | 1.4469 | 0.6430 | 0.6930 | 1.7906 | 1.7054 |
| | b | 1.2989 | 1.2451 | 2.2687 | 2.0794 | 1.2148 | 0.5318 | 0.6930 | 1.7871 | 1.5615 |
| | c | 1.2889 | 1.1104 | 1.6946 | 1.8068 | 0.6931 | 0.1837 | 0.6896 | 1.7538 | 1.3306 |
| | d | 1.1690 | 0.8972 | 1.6094 | 0.9957 | 0.2586 | 0.0843 | 0.6861 | 1.7401 | 1.2030 |
| **2** | a | 1.3924 | 1.2726 | Inf | Inf | 1.3321 | 0.6713 | 0.6930 | 1.7869 | 1.6699 |
| | b | 1.3603 | 1.1764 | 2.3026 | 1.7228 | 1.2969 | 0.6486 | 0.6930 | 1.7852 | 1.5363 |
| | c | 1.2894 | 1.1521 | 2.1972 | 1.4271 | 0.7510 | 0.2110 | 0.6896 | 1.7745 | 1.5150 |
| | d | 1.1825 | 0.8176 | 1.1820 | 0.7785 | 0.2049 | 0.0584 | 0.6896 | 1.7322 | 1.4793 |

Table 28 Continued

| Case | | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|---|
| | | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **3** | a | 1.4417 | 1.2643 | Inf | 2.9444 | 1.4412 | 0.6401 | 0.6930 | 1.7852 | 1.6756 |
| | b | 1.4286 | 1.2338 | Inf | 1.5686 | 1.1907 | 0.6194 | 0.6930 | 1.7888 | 1.4432 |
| | c | 1.1680 | 1.1756 | Inf | 1.3695 | 0.6057 | 0.1263 | 0.6896 | 1.7759 | 1.3792 |
| | d | 1.0685 | 0.9584 | 1.4854 | 0.9589 | 0.2286 | 0.0165 | 0.6815 | 1.7121 | 1.1976 |
| **4** | a | 1.4904 | 1.2363 | Inf | 2.3514 | 1.2062 | 0.6101 | 1.0101 | 1.0564 | 1.4588 |
| | b | 1.3569 | 1.1610 | Inf | 2.1203 | 0.8400 | 0.3662 | 1.0000 | 1.0137 | 1.4214 |
| | c | 1.2183 | 1.1750 | 2.2513 | 1.4553 | 0.6190 | 0.2470 | 0.9950 | 0.9989 | 1.4088 |
| | d | 1.1258 | 0.9087 | 1.5294 | 0.9163 | 0.2657 | 0.0484 | 0.9850 | 0.9589 | 1.2390 |
| **5** | a | 1.4178 | 1.2117 | Inf | 1.8971 | 1.1486 | 0.6553 | 0.6930 | 1.7602 | 1.7584 |
| | b | 1.3287 | 1.0830 | Inf | 1.2040 | 1.1415 | 0.5048 | 0.6930 | 1.7852 | 1.4845 |
| | c | 1.2870 | 1.1061 | Inf | 1.4321 | 0.6103 | 0.1756 | 0.6815 | 1.7307 | 1.6248 |
| | d | 1.2408 | 0.8691 | 1.7148 | 0.9330 | 0.2222 | 0.0498 | 0.6930 | 1.7836 | 1.2057 |
| **6** | a | 1.4432 | 1.2761 | Inf | Inf | 1.3782 | 0.6163 | 0.6896 | 1.7778 | 1.5904 |
| | b | 1.3670 | 1.2260 | Inf | 1.4816 | 1.2969 | 0.5346 | 0.6896 | 1.7591 | 1.4945 |
| | c | 1.3606 | 1.1859 | 2.0794 | 1.2404 | 0.6641 | 0.1656 | 0.6930 | 1.7852 | 1.4366 |
| | d | 1.1736 | 0.8586 | 1.0986 | 0.7199 | 0.1249 | 0.0017 | 0.6930 | 1.7370 | 1.4285 |
| **7** | a | 1.5781 | 1.3208 | Inf | 3.1355 | 1.3335 | 0.6209 | 0.6930 | 1.7888 | 1.6588 |
| | b | 1.3833 | 1.1581 | Inf | 2.3437 | 1.0906 | 0.5761 | 0.6919 | 1.7852 | 1.4432 |
| | c | 1.3582 | 1.0909 | Inf | 1.4028 | 0.5900 | 0.1148 | 0.6919 | 1.7800 | 1.3267 |
| | d | 1.2084 | 1.0532 | 1.4854 | 0.9500 | 0.2238 | 0.0144 | 0.6757 | 1.7070 | 1.1705 |
| **8** | a | 1.3917 | 1.2447 | Inf | 2.0369 | 1.2015 | 0.7126 | 0.6861 | 1.7215 | 1.5954 |
| | b | 1.3689 | 1.1804 | Inf | 1.7492 | 1.0849 | 0.5971 | 0.6896 | 1.7512 | 1.4808 |
| | c | 1.1770 | 1.1164 | 2.9957 | 1.6835 | 0.5213 | 0.1377 | 0.6896 | 1.7789 | 1.4257 |
| | d | 1.0899 | 0.8146 | 1.2730 | 0.7370 | 0.2114 | 0.0270 | 0.6919 | 1.7778 | 1.3404 |
| **9** | a | 1.4785 | 1.2537 | Inf | 2.3972 | 1.1342 | 0.6725 | 0.6861 | 1.6898 | 1.4772 |
| | b | 1.3550 | 1.1985 | Inf | 2.1979 | 0.8194 | 0.4164 | 0.6930 | 1.7646 | 1.4354 |
| | c | 1.1837 | 1.1628 | 2.3026 | 1.4718 | 0.6268 | 0.2648 | 0.6930 | 1.7778 | 1.3974 |
| | d | 1.1121 | 0.8957 | 1.5294 | 0.9163 | 0.2617 | 0.0562 | 0.6919 | 1.7855 | 1.3297 |
| **10** | a | 1.3736 | 1.1564 | Inf | 1.6094 | 1.1299 | 0.6594 | 0.6919 | 1.7501 | 1.6897 |
| | b | 1.2928 | 1.0386 | Inf | 1.4321 | 1.0402 | 0.4697 | 0.6919 | 1.7852 | 1.6895 |
| | c | 1.2401 | 1.0060 | Inf | 1.3291 | 0.6103 | 0.1611 | 0.6757 | 1.7147 | 1.3848 |
| | d | 1.2091 | 0.8745 | 1.7148 | 0.9330 | 0.2248 | 0.0424 | 0.6930 | 1.7819 | 1.2661 |

Again, we take Case 1 as an example, which shows the same results as the rest of the cases.

Like the conclusions drawn from simulation round I, with stricter parameter selection, the measurement results could increase reflecting levels of complexity. Based on the ranks resulting from the Six Sigma analysis, ApEn, SampEn and Maximal Lyapunov Exponent showed decreasing complexity for the four scenario of each case similar to the sigma level. SampeEn reflected significant results compared to ApEn and the Maximal Lyapunov Exponent, as illustrated in Figure 32, Figure 33 and Figure 34. Except for the significant rank based on SampEn, sometimes ApEn and Maximal Lyapunov Exponent also resulted in similar trends. Compared with ApEn and SampEn, results from PEn were pretty much the same, rather than decreasing, as shown in Figure 35. For the results based on SampEn, as shown in Table 24, we also concluded that the narrower tolerance of Sample Entropy still does not work well with a sample size of approximately 50.

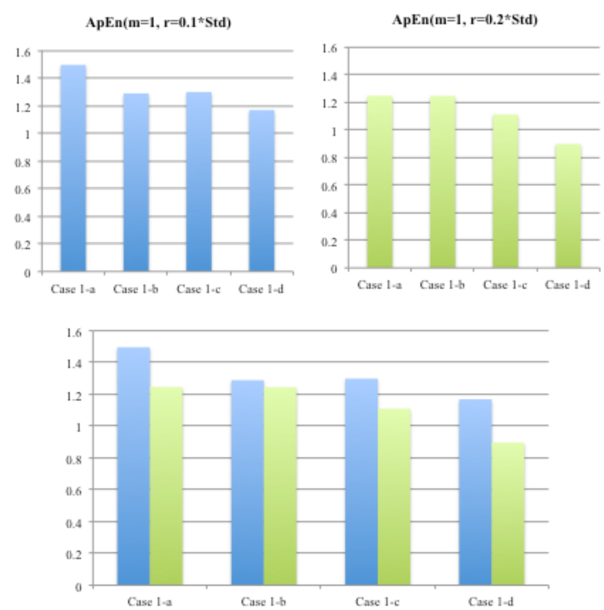



Figure 32  Complexity of Simulated Case 1 in Run II-A ranked by ApEn

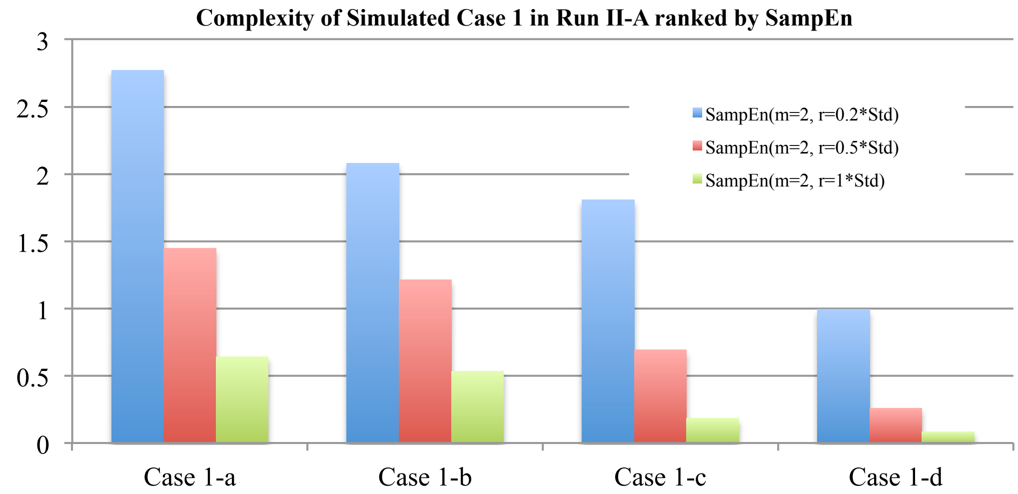



Figure 33 Complexity of Simulated Case 1 in Run II-A ranked by SampEn

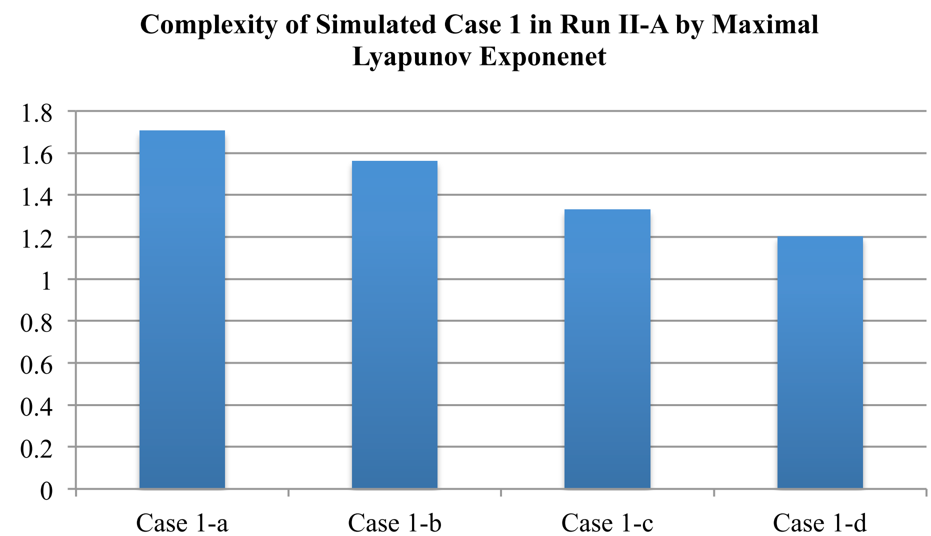



Figure 34 Complexity of Simulated Case 1 in Run II-A by Maximal Lyapunov Exponent

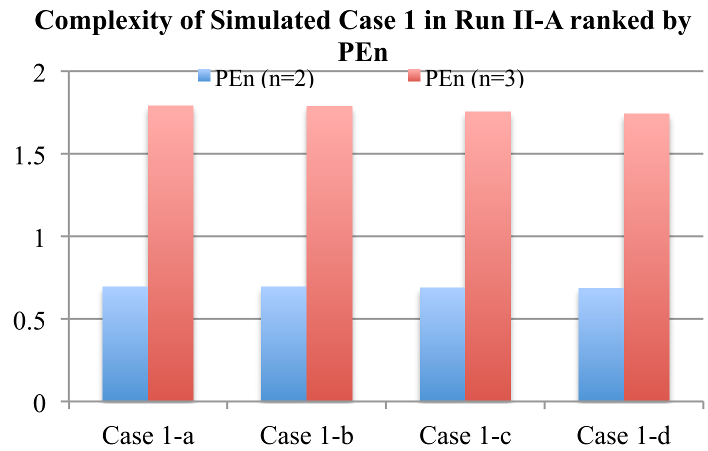



Figure 35 Complexity of Simulated Case 1 in Run II-A ranked by PEn

*Run II-B*

In Run II-B, additional random numbers were generated, based on the previous run for the set-up scenario of each case; this time each scenario had 110 PPC values to represent the different long-term performance scenarios of building construction. Similar to that which wwas conducted in the previous run, Six Sigma Analysis was first conducted to test the predicted complexity rank for the four scenarios of each case, as listed in Table 29. Notably, the simulated cases became less complex while the range of numbers shrank.

Table 29 Six Sigma for Simulated Cases Round II Run B (II-B)

| Case | Scenario a | Scenario b | Scenario c | Scenario d |
|---|---|---|---|---|
| Yield | 74.55% | 92.73% | 97.27% | 97.27% |
| Sigma Level | 3 | 3 | 4 | 4 |

The next step was to test the proposed entropy algorithms with the same parameter combinations and Maximal Lyapunov Exponent, then check the results for the different scenarios of the same case, and compare the results with the pretested rank. All the complexity calculation results are shown in Table 30. Based on the form, if we compare complexity results across the four different performance scenarios, the same conclusion were obtained as with the previous run. With ApEn, SampEn and Maximal Lyapunov Exponent, the complexity level decreased when the range of PPC values became narrower and narrower. Among them, the results from SampEn were more significant and clearer, as shown in Figure 36. The Pen method just achieved pretty close results for the four different scenarios; thus its results were not sensitive to different construction performance scenarios. Looking at the results from SampEn in more detail, we verified the first round simulation's results, which indicated that the smaller tolerance of SampEn cannot work well on a sample size of about 100--the larger tolerance choices for the algorithm work better.

Table 30 Complexity Measurement for Simulated Cases Round II Run B (II-B)

| **Case** | | **ApEn (m=1)** | | **SampEn (m=2)** | | | | **PEn** | | **Maximal Lyapunov Exponent** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **r=0.1*Std** | **r=0.2*Std** | **r=0.1*Std** | **r=0.2*Std** | **r=0.5*Std** | **r=Std** | **n=2** | **n=3** | |
| **1** | a | 1.7626 | 1.5907 | Inf | 2.3026 | 1.3289 | 0.6575 | 0.6928 | 1.7897 | 1.8824 |
| | b | 1.6714 | 1.5268 | 2.8332 | 1.8871 | 1.2944 | 0.58 | 0.6921 | 1.785 | 1.7395 |
| | c | 1.6345 | 1.4999 | 2.3418 | 2.0236 | 0.8651 | 0.2808 | 0.6931 | 1.7769 | 1.7367 |
| | d | 1.5802 | 1.1874 | 1.7834 | 1.1929 | 0.4436 | 0.0798 | 0.6779 | 1.7085 | 1.7107 |
| **2** | a | 1.6878 | 1.5764 | 3.0445 | 2.2687 | 1.2658 | 0.6261 | 0.6931 | 1.7907 | 1.8164 |
| | b | 1.6796 | 1.5667 | 2.6150 | 1.7975 | 1.1658 | 0.6070 | 0.6931 | 1.7912 | 1.7824 |
| | c | 1.6205 | 1.5315 | 1.7135 | 1.6614 | 0.8283 | 0.2963 | 0.6931 | 1.7815 | 1.7539 |
| | d | 1.4860 | 1.1161 | 1.7047 | 1.1161 | 0.3131 | 0.0606 | 0.6921 | 1.7703 | 1.6759 |

Table 30 Continued

| Case | | ApEn (m=1) r=0.1*Std | ApEn (m=1) r=0.2*Std | SampEn (m=2) r=0.1*Std | SampEn (m=2) r=0.2*Std | SampEn (m=2) r=0.5*Std | SampEn (m=2) r=Std | PEn n=2 | PEn n=3 | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|---|
| **3** | a | 1.7607 | 1.5874 | 3.0445 | 2.4953 | 1.3309 | 0.6298 | 0.6928 | 1.7817 | 1.7918 |
| | b | 1.7205 | 1.5163 | 2.9704 | 2.0971 | 1.1473 | 0.5768 | 0.6931 | 1.7897 | 1.7551 |
| | c | 1.6028 | 1.4398 | 2.6741 | 1.9924 | 0.8319 | 0.2438 | 0.6911 | 1.7845 | 1.7020 |
| | d | 1.5445 | 1.4052 | 2.1484 | 1.1701 | 0.4005 | 0.0665 | 0.6911 | 1.7794 | 1.6426 |
| **4** | a | 1.7835 | 1.6239 | Inf | 2.4510 | 1.4092 | 0.6638 | 0.6897 | 1.7490 | 1.7055 |
| | b | 1.7731 | 1.6144 | 3.0445 | 2.0794 | 1.1197 | 0.5042 | 0.6928 | 1.7730 | 1.6528 |
| | c | 1.5297 | 1.3713 | 2.3168 | 1.3903 | 0.7436 | 0.2572 | 0.6931 | 1.7758 | 1.5232 |
| | d | 1.4898 | 1.1531 | 1.8971 | 1.1073 | 0.4099 | 0.0514 | 0.6860 | 1.7697 | 1.4434 |
| **5** | a | 1.7423 | 1.6109 | 2.7300 | 2.0075 | 1.2866 | 0.6959 | 0.6911 | 1.7778 | 1.9147 |
| | b | 1.6054 | 1.4952 | 2.1203 | 1.6541 | 1.2109 | 0.5727 | 0.6911 | 1.7809 | 1.8253 |
| | c | 1.5665 | 1.4043 | 1.9818 | 1.5476 | 0.8231 | 0.2574 | 0.6897 | 1.7668 | 1.8151 |
| | d | 1.4946 | 1.1582 | 1.8681 | 1.1609 | 0.3198 | 0.0404 | 0.6921 | 1.7814 | 1.5867 |
| **6** | a | 1.7257 | 1.6138 | 3.0445 | 2.2095 | 1.2989 | 0.6929 | 0.6928 | 1.7877 | 1.8166 |
| | b | 1.7173 | 1.6058 | 2.5903 | 1.7778 | 1.2037 | 0.5167 | 0.6928 | 1.7871 | 1.7202 |
| | c | 1.6776 | 1.5588 | 1.7918 | 1.5079 | 0.7732 | 0.2633 | 0.6928 | 1.7800 | 1.6516 |
| | d | 1.4942 | 1.1323 | 1.6720 | 1.0841 | 0.2763 | 0.0228 | 0.6931 | 1.7685 | 1.6311 |
| **7** | a | 1.6960 | 1.6289 | Inf | 2.4277 | 1.5491 | 0.7611 | 0.6928 | 1.7806 | 1.7934 |
| | b | 1.6044 | 1.5960 | Inf | 1.9397 | 1.2977 | 0.6134 | 0.6931 | 1.7886 | 1.6120 |
| | c | 1.5425 | 1.4581 | 3.0204 | 1.8458 | 0.8248 | 0.2375 | 0.6921 | 1.7845 | 1.5795 |
| | d | 1.4528 | 1.2139 | 2.1484 | 1.1676 | 0.3988 | 0.0660 | 0.6897 | 1.7789 | 1.5488 |
| **8** | a | 1.7516 | 1.6311 | Inf | 2.1972 | 1.2550 | 0.6473 | 0.6897 | 1.7474 | 1.8293 |
| | b | 1.7031 | 1.6005 | Inf | 1.8788 | 1.1576 | 0.6253 | 0.6921 | 1.7814 | 1.8229 |
| | c | 1.6120 | 1.4907 | 2.3749 | 1.8211 | 0.7644 | 0.2299 | 0.6928 | 1.7887 | 1.7569 |
| | d | 1.4399 | 1.1118 | 1.8632 | 1.1345 | 0.3425 | 0.0447 | 0.6930 | 1.7811 | 1.6791 |
| **9** | a | 1.7794 | 1.6204 | Inf | 2.5177 | 1.3528 | 0.6912 | 0.6897 | 1.7490 | 1.8270 |
| | b | 1.7929 | 1.6086 | 2.9444 | 2.0281 | 1.1285 | 0.5274 | 0.6928 | 1.7730 | 1.7377 |
| | c | 1.5419 | 1.4055 | 2.3308 | 1.4110 | 0.7411 | 0.2577 | 0.6931 | 1.7758 | 1.6250 |
| | d | 1.4755 | 1.1362 | 1.9136 | 1.1094 | 0.4079 | 0.0499 | 0.6860 | 1.7697 | 1.5832 |
| **10** | a | 1.7216 | 1.5845 | 3.0445 | 1.9021 | 1.2590 | 0.6945 | 0.6921 | 1.7800 | 1.9351 |
| | b | 1.5910 | 1.4775 | 2.0149 | 1.6701 | 1.1959 | 0.5600 | 0.6897 | 1.7767 | 1.8273 |
| | c | 1.5568 | 1.3987 | 1.9459 | 1.6422 | 0.8260 | 0.2553 | 0.6880 | 1.7607 | 1.8151 |
| | d | 1.4822 | 1.1474 | 1.8681 | 1.1609 | 0.3202 | 0.0397 | 0.6928 | 1.7812 | 1.5180 |

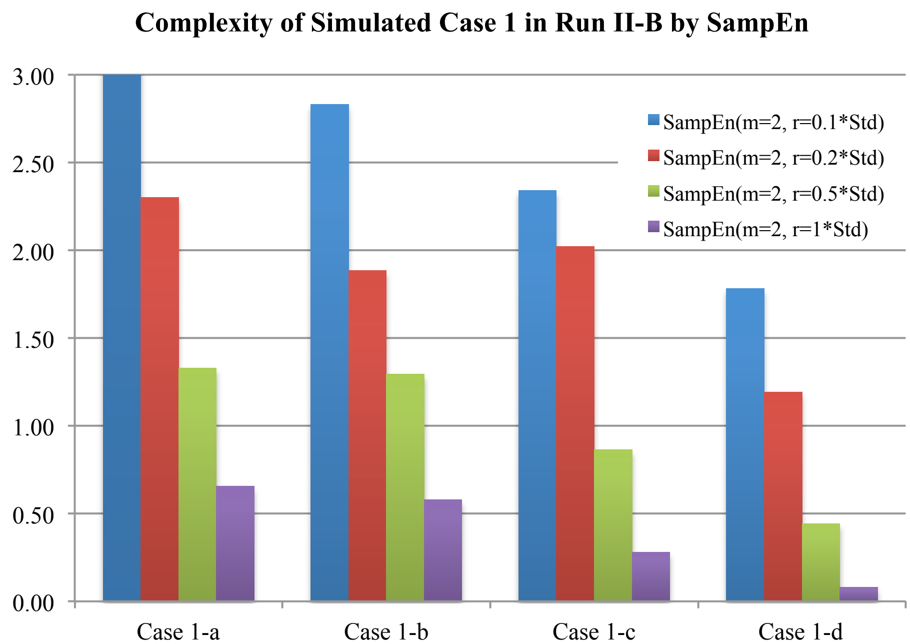



Figure 36 Complexity of Simulated Case 1 in Run II-B by SampEn

*Run II-C*

In the last run of the Round II simulation, Run II-C, 100 more random numbers were generated for the four different scenarios of each pilot case; each case had 210 PPC values in each of its performance scenarios. Six Sigma Analysis was conducted and again found decreased complexity from scenario a to scenario d. Proposed entropy algorithms with different parameter combinations and Maximal Lyapunov Exponent were tested in Run II-C again. Their results are shown in Table 31. For each specific entropy model, stricter parameter selections resulted in higher complexity levels for the same case, as easily seen from the results.

Table 31 Complexity Measurement for Simulated Cases Round II Run C (II-C)

| **Case** | | **ApEn (m=1)** | | **SampEn (m=2)** | | | | **PEn** | | **Maximal Lyapunov Exponent** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **r=0.1*Std** | **r=0.2*Std** | **r=0.1*Std** | **r=0.2*Std** | **r=0.5*Std** | **r=Std** | **n=2** | **n=3** | |
| **1** | a | 1.9857 | 1.9102 | 2.7881 | 1.9755 | 1.3129 | 0.6542 | 0.6930 | 1.7886 | 2.0514 |
| | b | 1.9371 | 1.8103 | 2.4696 | 1.7253 | 1.0176 | 0.4120 | 0.6931 | 1.7866 | 2.0937 |
| | c | 1.6871 | 1.3506 | 1.6924 | 1.1964 | 0.4168 | 0.1225 | 0.6931 | 1.7888 | 2.0222 |
| | d | 1.5318 | 1.2400 | 1.2577 | 0.7159 | 0.2283 | 0.0500 | 0.6918 | 1.7741 | 1.9524 |
| **2** | a | 2.0315 | 1.9442 | 2.8744 | 2.3775 | 1.1917 | 0.6978 | 0.6929 | 1.7879 | 2.0235 |
| | b | 1.9882 | 1.9323 | 2.5953 | 2.0116 | 1.2470 | 0.5039 | 0.6930 | 1.7906 | 2.0227 |
| | c | 1.8907 | 1.7453 | 2.5390 | 1.9538 | 0.9729 | 0.3960 | 0.6931 | 1.7865 | 2.0518 |
| | d | 1.8373 | 1.4227 | 2.3265 | 1.4556 | 0.5521 | 0.1222 | 0.6912 | 1.7854 | 2.0002 |
| **3** | a | 1.9848 | 1.9468 | 2.9704 | 2.4496 | 1.2551 | 0.6445 | 0.6929 | 1.7856 | 2.0456 |
| | b | 1.9672 | 1.8491 | 2.5777 | 2.0412 | 1.1930 | 0.5709 | 0.6922 | 1.7736 | 2.0193 |
| | c | 1.9495 | 1.8112 | 2.4159 | 1.9072 | 0.9743 | 0.3900 | 0.6931 | 1.7889 | 1.9855 |
| | d | 1.7720 | 1.4130 | 2.1193 | 1.3574 | 0.5528 | 0.1215 | 0.6926 | 1.7878 | 1.8800 |
| **4** | a | 1.9875 | 1.9646 | 2.9575 | 2.2156 | 1.2851 | 0.6922 | 1.7824 | 1.7886 | 1.9956 |
| | b | 1.9736 | 1.9374 | 2.7150 | 1.9524 | 1.0595 | 0.6930 | 1.7848 | 1.7866 | 1.8824 |
| | c | 1.9554 | 1.7755 | 2.6148 | 1.7918 | 0.9437 | 0.6906 | 1.7690 | 1.7888 | 1.4863 |
| | d | 1.8871 | 1.4310 | 2.5275 | 1.4234 | 0.5841 | 0.6926 | 1.7890 | 1.7741 | 1.2835 |
| **5** | a | 2.0532 | 1.9765 | 3.0819 | 2.2502 | 1.3080 | 0.6994 | 0.6922 | 1.7888 | 2.3077 |
| | b | 2.0099 | 1.9088 | 2.7344 | 2.1484 | 1.2780 | 0.6191 | 0.6929 | 1.7900 | 2.0792 |
| | c | 1.9460 | 1.8376 | 2.4159 | 2.0401 | 1.0343 | 0.4321 | 0.6918 | 1.7717 | 1.9987 |
| | d | 1.8694 | 1.4069 | 2.1041 | 1.4400 | 0.5689 | 0.1293 | 0.6931 | 1.7880 | 1.9921 |
| **6** | a | 2.0383 | 1.9568 | 2.8622 | 2.4371 | 1.1846 | 0.5871 | 0.6922 | 1.7858 | 1.9277 |
| | b | 2.0078 | 1.8569 | 2.8163 | 2.0254 | 1.2496 | 0.5983 | 0.6926 | 1.7895 | 1.7890 |
| | c | 1.9916 | 1.7562 | 2.5802 | 1.9238 | 0.9532 | 0.3758 | 0.6930 | 1.7858 | 1.5620 |
| | d | 1.8407 | 1.4233 | 2.3119 | 1.4266 | 0.5315 | 0.1015 | 0.6922 | 1.7882 | 1.3290 |
| **7** | a | 1.9857 | 1.9102 | 3.2055 | 2.1302 | 1.3859 | 0.6674 | 0.6922 | 1.7837 | 2.0556 |
| | b | 1.9371 | 1.8103 | 3.0910 | 1.9763 | 1.2562 | 0.6467 | 0.6930 | 1.7853 | 2.0522 |
| | c | 1.6871 | 1.3506 | 2.4472 | 1.9132 | 0.9163 | 0.3622 | 0.6931 | 1.7904 | 1.9851 |
| | d | 1.5318 | 1.0400 | 2.0794 | 1.4001 | 0.5477 | 0.1023 | 0.6929 | 1.7823 | 1.9258 |
| **8** | a | 1.9857 | 1.9102 | 3.2055 | 2.1302 | 1.3859 | 0.6674 | 0.6922 | 1.7837 | 2.0556 |
| | b | 1.9371 | 1.8103 | 3.0910 | 1.9763 | 1.2562 | 0.6467 | 0.6930 | 1.7853 | 2.0522 |
| | c | 1.6871 | 1.3506 | 2.4472 | 1.9132 | 0.9163 | 0.3622 | 0.6931 | 1.7904 | 1.9851 |
| | d | 1.5318 | 1.0400 | 2.0794 | 1.4001 | 0.5477 | 0.1023 | 0.6929 | 1.7823 | 1.9258 |

Table 31 Continued

| Case | | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|---|
| | | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| **9** | a | 1.9965 | 1.9634 | 2.8904 | 2.1418 | 1.2741 | 0.6622 | 0.6922 | 1.7824 | 2.1186 |
| | b | 1.9848 | 1.9570 | 2.5903 | 1.9525 | 1.0642 | 0.5557 | 0.6930 | 1.7848 | 2.0124 |
| | c | 1.9491 | 1.7821 | 2.4999 | 1.7897 | 0.9450 | 0.4015 | 0.6906 | 1.7690 | 2.0055 |
| | d | 1.8645 | 1.4279 | 2.3945 | 1.4266 | 0.5848 | 0.1237 | 0.6926 | 1.7890 | 1.8763 |
| **10** | a | 2.0182 | 1.9698 | 3.0910 | 2.1471 | 1.3016 | 0.6921 | 0.6926 | 1.7895 | 2.2940 |
| | b | 1.9843 | 1.9004 | 2.6027 | 2.0331 | 1.2660 | 0.6092 | 0.6926 | 1.7895 | 2.1028 |
| | c | 1.9280 | 1.8095 | 2.5337 | 1.9517 | 1.0363 | 0.4307 | 0.6912 | 1.7707 | 2.0149 |
| | d | 1.8715 | 1.4000 | 2.1041 | 1.4400 | 0.5689 | 0.1288 | 0.6930 | 1.7861 | 2.0051 |

For ApEn and Maximal Lyapunov Exponent, the complexity measurement decreased for the four scenarios of the same case. However, these results were not as significant as the results for SampEn as presented in Figure 37. Permutation Entropy's results were roughly the same for the different performance scenarios. When the compared template contained two (2) units, the results were around 0.69, and they were approximately 1.78 when three (3) units were compared each time. For Sample Entropy, not only did it offer significant results for different construction performance scenarios, both its narrow and wide tolerances worked for the larger sample size of around 200.

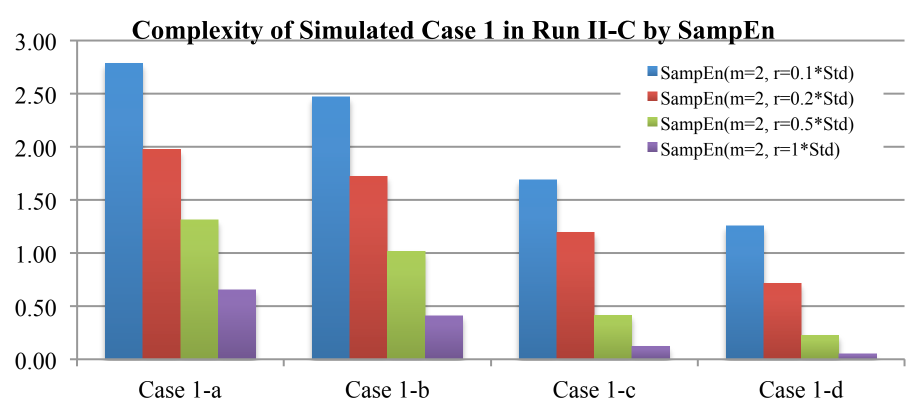



Figure 37 Complexity of Simulated Case 1 in Run II-C by SampEn

### 6.2.3 Normalization of Complexity Measurements in Round II Simulation

As in Round I of the pilot case analysis and the comparative complexity measurement analyses, the Round II simulation results were normalized in order to compare different entropy complexity results. To normalize the measurements, we divided the maximal measurement from each group of results. In order to verify the previous comparative analysis for the complexity measurements in simulation Round II, Case 1 was selected to showcase the normalization procedure and its results. The normalization procedure for the rest of the cases was the same, but those cases are not mentioned in this dissertation.

For the first run of the Round II simulation, the normalized complexity measurements for the four different performance circumstances of Case 1 are summarized in Table 32, and then plotted in Figure 38. All the normalized results of ApEn, SampEn and Maximal Lyapunov Exponent follow the pretested performance ranks created by the Six Sigma Analysis. Except for the fact that the small tolerance of SampEn did not work well for sample size 60, the other results of SampEn were more significant than the results of the other methods. Compared with ApEn, SampEn and Maximal Lyapunov Exponent, the results from PEn were pretty much the same, rather than showing a decreasing track, as shown in Figure 38.

Table 32 Normalized Complexity Measurements for Case 1 in Simulation Run II-A

| **Case** | | **ApEn (m=1)** | | **SampEn (m=2)** | | | | **PEn** | | **Maximal Lyapunov Exponent** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **r=0.1*Std** | **r=0.2*Std** | **r=0.1*Std** | **r=0.2*Std** | **r=0.5*Std** | **r=Std** | **n=2** | **n=3** | |
| **1** | a | 1.0000 | 1.0000 | N/A | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | b | 0.8684 | 0.9989 | N/A | 0.7500 | 0.8396 | 0.8271 | 1.0000 | 0.9980 | 0.9156 |
| | c | 0.8617 | 0.8908 | N/A | 0.6517 | 0.4790 | 0.2857 | 0.9951 | 0.9794 | 0.7802 |
| | d | 0.7816 | 0.7198 | N/A | 0.3591 | 0.1787 | 0.1311 | 0.9900 | 0.9718 | 0.7054 |

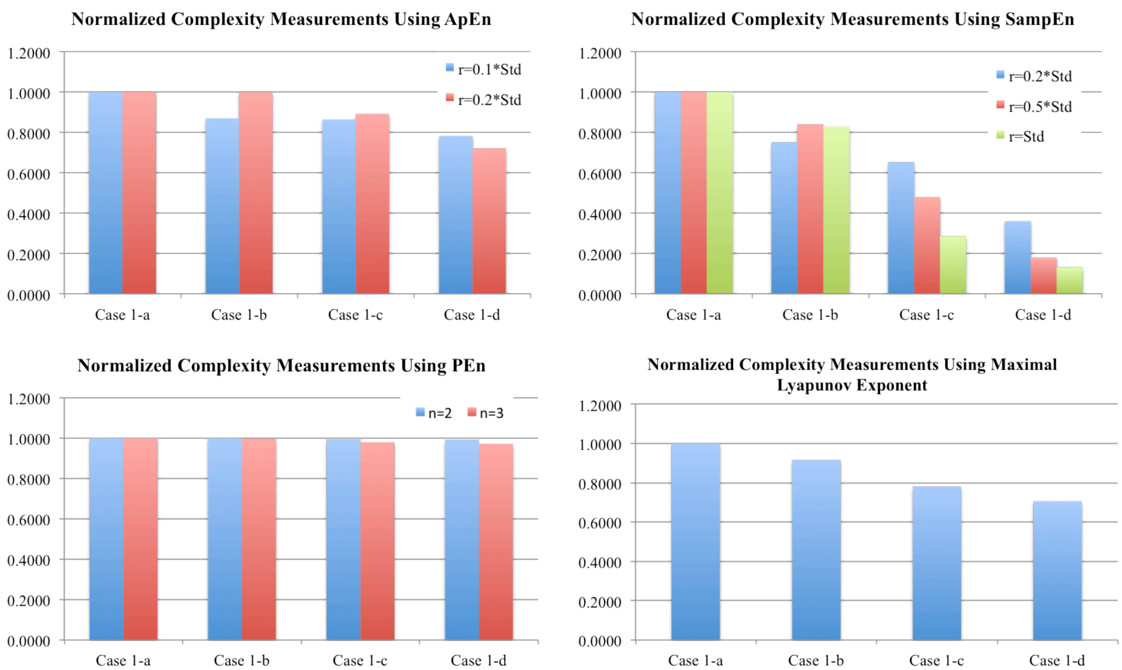



Figure 38 Normalized Complexity Measurements for Four Different Performance Scenarios of Case 1 in Simulation Run II-A

For Run B of the Round II simulation, the normalized complexity measurements for the four different performance circumstances of Case 1 are summarized in Table 33, and then plotted in Figure 39. Based on the normalized results, and comparing the complexity results across the four different performance scenarios, we obtained the same conclusions as those from Run II-B's descriptive analysis. With ApEn, SampEn and the Maximal Lyapunov Exponent, the complexity level decreased when the range of PPC values became narrower and narrower. Among them, the results from SampEn were more significant and clearer compared with the others. PEn obtained pretty close results for the four different scenarios, meaning that its results wewre not sensitive to different construction performance scenarios. In the end, the smallest tolerance of SampEn still did not work well for a sample size of around 100.

Table 33 Normalized Complexity Measurements for Case 1 in Simulation Run II-B

| Case | | ApEn (m=1) | | SampEn (m=2) | | | | PEn | | Maximal Lyapunov Exponent |
|---|---|---|---|---|---|---|---|---|---|---|
| | | r=0.1*Std | r=0.2*Std | r=0.1*Std | r=0.2*Std | r=0.5*Std | r=Std | n=2 | n=3 | |
| 1 | a | 1.0000 | 1.0000 | N/A | 1.0000 | 1.0000 | 1.0000 | 0.9996 | 1.0000 | 1.0000 |
| | b | 0.9483 | 0.9598 | N/A | 0.8196 | 0.9740 | 0.8821 | 0.9986 | 0.9974 | 0.9241 |
| | c | 0.9273 | 0.9429 | N/A | 0.8788 | 0.6510 | 0.4271 | 1.0000 | 0.9928 | 0.9226 |
| | d | 0.8965 | 0.7465 | N/A | 0.5181 | 0.3338 | 0.1214 | 0.9781 | 0.9546 | 0.9088 |

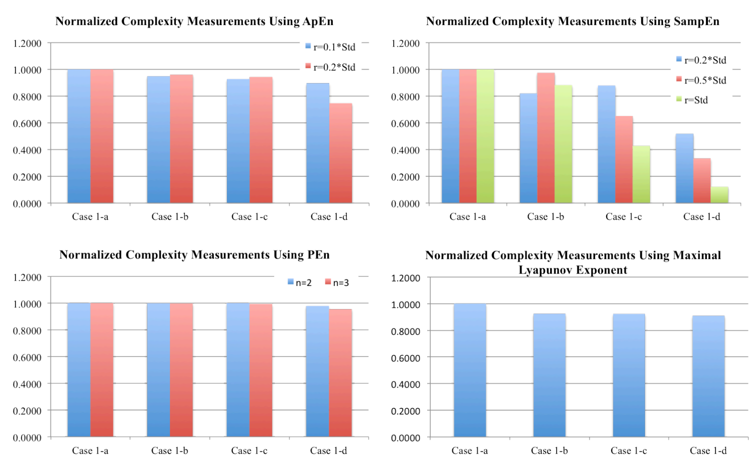



Figure 39 Normalized Complexity Measurements for Four Different Performance Scenarios of Case 1 in Simulation Run II-B

In the end, the normalized complexity measurements for the four different performance circumstances of Case 1 in the last run of simulation Round II, which have the largest sample sizes, are summarized in Table 34, and Figure 40. For ApEn and Maximal Lyapunov Exponent, the normalized results showed decreases in complexity measurements for the four scenarios of the same case; however, their results were not as significant as the results of SampEn were. Permutation Entropy's results were roughly the same for the different performance scenarios. For Sample Entropy, both its narrow and wide tolerances worked for the larger sample size of around 200.

Table 34 Normalized Complexity Measurements for Case 1 in Simulation Run II-C

| **Case** | | **ApEn (m=1)** | | **SampEn (m=2)** | | | | **PEn** | | **Maximal Lyapunov Exponent** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **r=0.1*Std** | **r=0.2*Std** | **r=0.1*Std** | **r=0.2*Std** | **r=0.5*Std** | **r=Std** | **n=2** | **n=3** | |
| **1** | a | 1.0000 | 1.0000 | N/A | 1.0000 | 1.0000 | 1.0000 | 0.9996 | 1.0000 | 1.0000 |
| | b | 0.9483 | 0.9598 | N/A | 0.8196 | 0.9740 | 0.8821 | 0.9986 | 0.9974 | 0.9241 |
| | c | 0.9273 | 0.9429 | N/A | 0.8788 | 0.6510 | 0.4271 | 1.0000 | 0.9928 | 0.9226 |
| | d | 0.8965 | 0.7465 | N/A | 0.5181 | 0.3338 | 0.1214 | 0.9781 | 0.9546 | 0.9088 |

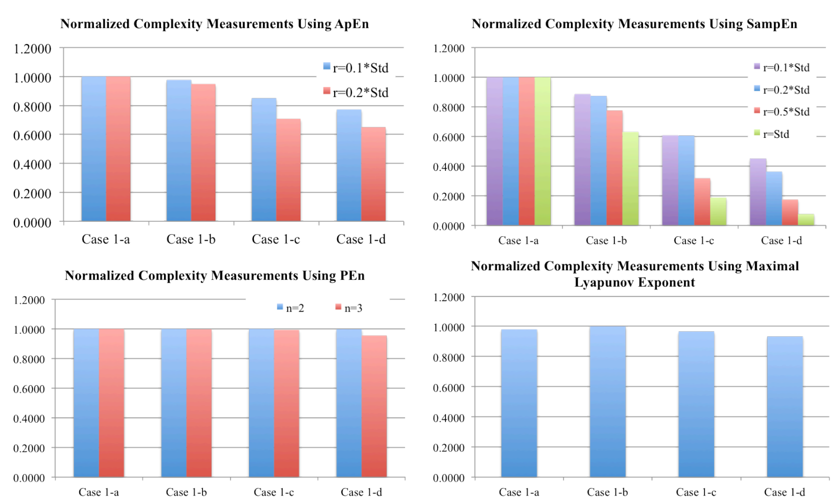



Figure 40 Normalized Complexity Measurements for Four Different Performance Scenarios of Case 1 in Simulation Run II-C

### 6.2.4 Inferential Analysis of Comparative Analysis for Round II

The purpose of the Round II simulations with three different runs that included three different sample sizes and four different scenarios for each case was to test the workability of the proposed entropy algorithms for different construction performances. Both the descriptive analysis of the complexity measurements and the analysis of normalized results provided the same results. As found in the basic Six Sigma Analysis of the four different scenarios, the wider the range of the random PPC values, the more complex the construction performances were. The ranks of each proposed entropy algorithm are summarized in Table 35.

The Sigma Level of each case provided a preliminary performance rank of the four different scenarios. The complexity rank based on both ApEn and SampEn showed that the lower the sigma level, the higher the complexity level. This rank was also verified by the Maximal Lyapunov Exponent. Based on the different sample sizes, we also concluded that only Sample Entropy worked well for describing different construction scenarios based on different sample sizes. The larger the sample size, the narrower the tolerance was for Sample Entropy. Approximate Entropy and the Maximal Lyapunov Exponent provided the same complexity rank for different performance scenarios with small samples. For large samples, neither of them was sensitive enough to differentiate the performance levels.

Table 35 Proposed Methods' Workability of Dealing with Different Performance Scenarios

| Proposed Algorithms | Workability for Different Performance Scenarios | | |
|---|---|---|---|
| | 60 | 110 | 210 |
| ApEn with Small Tolerance | Yes | No | No |
| SampEn with Small Tolerance | No | Yes | Yes |
| SampEn with Large Tolerance | Yes | Yes | No |
| PEn | No | No | No |
| Maximal Lyapunov Exponent | Yes | Yes | No |
| Six Sigma Analysis | Yes | Yes | Yes |

### *6.3 Summary of Simulation and Comparative Analysis*

As discussed in the previous sections of this chapter, two rounds of simulation were conducted based on the ten (10) pilot cases in order to create different cases with different construction performance patterns. The simulation procedures expanded the scenarios we used to test the proposed entropy algorithms by creating samples with different sizes and different ranges of values. Applying these simulated cases in further comparative analysis for selected entropy algorithms provided enough evidence to draw solid conclusion and to make reasonable selections.

MATLAB was used to generate more random PPC numbers and simulate different construction performance scenarios for both rounds of simulation based on the pilot cases. The first round of simulations tested the workability of proposed entropy algorithms and their parameter selections for different sample sizes. For this reason, 50, 100 and 200 additional random PPC values that followed the same probability distribution of each pilot case were generated separately for the three runs embedded in the first round simulation. The second round simulation added one more dimension, the different performance of each case, to test the workability of each proposed entropy algorithm. In this round, 50, 100 and 200 additional PPC values were generated again for each pilot case; however, this time we set four different performance ranges to each simulated case. They were (Mean-3*Standard Deviation, Mean-3*Standard Deviation) for scenario a of each case, (Mean-1.5*Standard Deviation, Mean-1.5*Standard Deviation) for scenario b, (Mean-0.5*Standard Deviation, Mean-0.5*Standard Deviation) for scenario c and (Mean-0.25*Standard Deviation, Mean-0.25*Standard Deviation) for scenario d. In each run, each pilot case had four simulated scenarios with specific PPC ranges. The ability to separate different performance circumstances for each entropy model was used in this round.

Approximate Entropy, Sample Entropy and Permutation Entropy, each with several different parameter choices, were tested in two rounds of simulation, for a total of six runs. Six Sigma Analysis and Maximal Lyapunov Exponent were also used in the comparative analysis to provide more references and points to check the results of the entropy algorithms. The calculated results in each simulated run were presented in the tables discussed in the previous sections of this chapter. Their normalized results also double checked the comparative analysis results. Based on those results, a summary of their workability to deal with different sample sizes and different performance circumstances is presented in Table 36.

Table 36 Workability of Proposed Entropy Algorithms to Deal with Different Size and Performance

| **Proposed Algorithms** | **Workability** | | |
|---|---|---|---|
| | **Different Sample Size** | | **Different Performance** |
| | **Small** | **Large** | **(Different PPC Range)** |
| ApEn with Small Tolerance | Yes | No | Yes |
| ApEn with Small Tolerance | Yes | No | Yes |
| SampEn with Small Tolerance | No | Yes | Yes |
| SampEn with Large Tolerance | Yes | No | No |
| PEn | No | No | No |
| Maximal Lyapunov Exponent | Yes | No | Yes |
| Six Sigma Analysis | No | No | Yes |

According to the workability summary, Approximate Entropy works well with small sample sizes. When the sample size was small (less than 200), it shows both the size and increasing complexity, which we see in the pilot case and in the first and second simulation runs. However, this result becomes close regarding different performance circumstances when the sample size exceeds 200. For Sample Entropy, by choosing a different tolerance parameter, both small and large sample sizes reflect the complexity level of different construction performance scenarios. Small tolerance provides significant results for different performance scenarios for a large sample, while large tolerance works better for a small sample size. The final entropy algorithm tested, Permutation Entropy, cannot work in either simulated scenario, as its results are based on the fluctuation dimension only, rather than value of each number.

The added comparison with Six Sigma Level and Maximal Lyapunov Exponent for each simulation case also provided several more conclusions to ponder. As long as the performance follows the same distribution pattern, Six Sigma Level is not a good choice for performance evaluation since it is a statistical value for each case, especially with increasing sample size. However, with different performance scenarios, a Sigma Level for each case provides a pretested rank of different scenarios to compare with the entropy results. The Maximal Lyapunov Exponent's performance has not been widely tested in the construction management field yet, but we tested in the comparative analysis. From the test results, we see that it could also tell the complexity/or dynamic level of different performance circumstances for small sample sizes. The complexity rank of its results is similar to the rank based on Sample Entropy; however, its results are not as significant as those for Sample Entropy.

# CHAPTER VII

# SUMMARY AND CONCLUSIONS

This study has been a pioneering complexity study, especially quantitative complexity measurement, in building construction systems. This is the first study to apply entropy algorithms for complexity measurement in building construction from the perspective of chaos theory. This final chapter includes a summary of the research, findings and conclusions, limitations of the study, contributions and recommendations, and future research directions.

## *7.1 Summary of the Research*

According to the abundant literature search, research in construction has already started to probe the complex nature of construction operation and considers it a fundamental part of construction management. However, none of the existing studies provide a scientific and comprehensive analysis of this complexity, an inclusive understanding of the complexity of construction, and the most important, a transmittable and understandable model to calculate the complexity of building construction. As a result, the overall objective of this research was to create a novel mathematical model to calculate complexity quantitatively in order to fill the research gap of a missing mathematical model for the study of complexity in building construction.

In support of the overall objective, there were three supporting specific objectives that were explored, including: a comprehensive and embedded theoretical analysis for complexity in building construction; appropriate entropy models to quantitatively present complexity in building construction; and measurement of complexity in building construction cases using the most efficient entropy algorithm.

Organizationally, the study began with a review of literature concerning complexity research in building and construction domains and other related subjects, chaos theory, and entropy algorithms. In conjunction with the literature review, the first hypothesis of this study was proposed, as there is no inclusive understanding of complexity in building construction systems, regarding construction's real nature--whether it is mainly characterized as complex or just complicated. In order to answer this question, a comprehensive theoretical framework of complexity study was built to reflect the complex nature of building construction, and to serve as the theoretical foundation. Chaos theory was treated as a companion to complexity and representation of a more complex order to link building construction systems, complexity studies, and measurement together. From the perspective of chaos theory, an entropy-measuring model linked with the complex nature in building construction was proposed to provide a scientific and quantitative analysis of complexity in building construction, which constituted the second hypothesis of this study. Together, the literature review and theoretical analysis established a baseline for complexity measurement in building construction;  three (3) widely used entropy algorithms were selected for the comparative analysis.

The comparative analysis was the other indispensable characteristic of this study in order to fulfill the research goal to find the most appropriate and efficient entropy algorithm for complexity measurement in building construction. For the pilot comparative study of proposed entropy algorithms, ten (10) pilot cases were randomly selected from the twenty-two (22) real construction cases that Fernandez-Solis used for his study of chaos theory in building construction (Fernandez-Solis, 2013). This provided the initial results, along with some critical criteria for further analysis. Two rounds of simulations based on the pilot cases were used to explore both different sample sizes and construction performance scenarios for further comparison of entropy algorithms, as well as Maximal Lyapunov Exponent and Six Sigma Analysis. Simulations were conducted by MATLAB to generate more random numbers based on the pilot cases for the purpose of representing different circumstances in building construction performance associated with different sample sizes. Approximate Entropy, Sample Entropy and Permutation Entropy, each with several different parameter choices, were tested in two rounds of simulation, for a total of six runs. Six Sigma Analysis and Maximal Lyapunov Exponent were also used in the comparative analysis to provide more references and checkpoints for the results of the entropy algorithms.

The complexity measurement results from the proposed methods were then analyzed in a descriptive perspective and an inferential manner. The results were also verified by the analysis of their normalized complexity measurements. The outcomes of those analyses have been compiled and established in the findings and conclusions of the study. The rigor of the descriptive and inferential analyses were conveyed in Chapter VI.

## *7.2 Findings and Conclusions*

This study reviewed and summarized the basic concepts of complexity and connected it with building construction systems. If building construction's project delivery system can be loosely categorized as a production system, rather than a complicated system that purely combines intricate components, it is a proven interdependent, open, emergent, non-linear and adaptive complex system. The output of a construction organization and production system is not proportional to the inputs, and the final product—a finished building--is not only the simple sum of its material and partial components. With a developing understanding of the complex nature of building construction and the evolutionary mechanism of the system, there is a new era for complex system research and complex sciences in construction research and practice.

From the perspective of chaos theory, an entropy-measuring model linked with the complex nature of building construction may provide a scientific and quantitative measurement of complexity in building construction systems. Among the three proposed commonly used entropy algorithms, which are Approximate Entropy, Sample Entropy and Permutation Entropy, Sample Entropy was the most efficient method for complexity measurement in building construction, using PPC values based on the comparative analysis of pilot cases and simulated cases.

The three proposed entropy models are not simple calculations. Each of them provides different sensitivities by using different values for specific parameters embedded within the algorithms. The choice of tolerance, also named effective filtering level, is the most important determinant for the sensitivity of Approximate Entropy and Sample Entropy; while the order, which is the number of units in the compared template, decides the sensitivity of Permutation Entropy. Based on the calculated results, it could be concluded that the narrower the tolerance, the more sensitive are the Approximate Entropy and Sample Entropy algorithms. That is to say, for the same case, it would have a higher complexity measurement. The same sensitivity increase could be observed for Permutation Entropy when its order becomes larger.

Permutation Entropy tracks complexity based simply on the shape of time series data without the magnitude of the shape. Thus, it is not workable for building construction performance complexity measurements based on PPC values. Approximate Entropy is only sensitive to different performance scenarios for small sample sizes. That is to say, it is useful for small construction projects or short term performance complexity measurements. Both its workability and sensitivity markedly decreases when the sample size exceeds two-hundred (200). This restricts the performance of Approximate Entropy for complexity measurement of large and long-term building construction projects.

Compared with Approximate Entropy and Permutation Entropy, the characteristics of Sample Entropy offset the defects of these two methods. It worked sensitively and efficiently to tell different construction performance circumstances apart through significant complexity measurements for either small or large sample sizes. For small sample sizes (10-100), which could refer to short-term construction systems, by choosing a larger tolerance, it provides significant measurement for different performance scenarios. And for large sample sizes, which means long-term construction systems, it maintains the same workability and significance level by decreasing the value of its tolerance. The other important reason to select Sample Entropy for complexity measurement in building construction, rather than Approximate Entropy and Permutation Entropy, is because the rank of different performance scenarios based on Sample Entropy is similar to the rank based on Six Sigma Level or the Maximal Lyapunov Exponent, if they were applicable for the comparison.

### *7.3 Limitations of the Study*

As a pioneering study of a quantitative complexity measurement in building construction systema, this study was geared towards finding the most efficient entropy algorithms for complexity in building construction by using PPC values based on the theoretical and comparative analysis of three commonly used entropy algorithms. With this objective, the study endured limiting elements through the course of the research.

The study encountered three main limitations. The first limitation was associated with the values used to calculate complexity in building construction, which is weekly PPC data records introduced by lean construction. Even though it represents the overall reliability of production planning and workflow, rather than focusing on a single perspective, such as schedule and cost, it has not been widely used in the construction industry as a main source for data records. However, it is still probable that other factors not addressed in this study could be used to calculate complexity in building construction. The second limitation was obtaining a large enough sample to generate conclusions about the study that could be generalized. Because of the application of PPC records from lean construction is still in its early stage in building construction, only a few PPC record cases exist in the previous data record. This restricts the available cases for use in the comparative analysis. This study used Monte Carlo simulation to generate more random numbers in order to represent different construction performance scenarios; but it is still possible that some specific cases we did not cover could have significant effects on the final conclusions. The third limitation was associated with the proposed entropy algorithms for further comparative analysis. Approximate Entropy, Sample Entropy and Permutation Entropy are three of the most commonly used entropy algorithms in existing research of complexity measurement based on the literature review, but they are not the only entropy algorithms for complexity measurement. Due to limitation of effort and scope, we could not compare all the entropy models in this study. However, we cannot rule out the probability that there are other entropy algorithms that may work better for complexity measurement. The study of other possible entropy algorithms for complexity measurement in building construction would be one direction for future research.

A multitude of factors contributed to the results and outcomes of the research and the study recognizes and acknowledges that. Because of this, it is virtually impossible to account for every factor or contributing nuance and it is beyond the intended scope of this study.

### *7.4 Contributions and Recommendations*

It is the intent of every dissertation to put forth a measurable and impactful piece of knowledge with the hope that it will become a placeholder in the framework of the given field of study. Within the defined range of study, this research has provided valid and meaningful results through a comprehensive theoretical and mathematical process.

First of all, the comprehensive theoretical framework of complexity research in building construction, based on previous literature and the conclusion that building construction could be considered a complex system, provided a revolutionary way to understand all the possible uncertainties and order changes during the construction process. This study, along with other pioneers' work, is just the beginning of complexity research in building construction systems. There is much left to accomplish in this field, but this study provides a solid foundation for future research on complexity in building construction systems.

The contributions of this study also include an effort to fill a knowledge gap and by finding a quantitative measurement of complexity in building construction. The complexity measurement based on Sample Entropy could provide a more objective and direct understanding of the complex nature of building construction. Compared with the previous qualitative analyses, Sample Entropy could avoid the subjectivity of evaluators and set a unified standard for complexity measurement in building construction, especially for comparing the complexity level of different construction performance circumstances, and to track construction performance from the complexity perspective, and even to test approaches in order to reduce the complexity level and to improve construction performance. All of this can be further studied in future research.

Last but not the least, in order to use Sample Entropy to measure complexity in real construction projects, several recommendations should be considered. Approximate tolerance, or so called effective filtering level, should be selected in order to reach the desired sensitivity. An effective filtering level of 0.5 or 1.0 times the standard deviation of the recorded PPC values could be used for cases with 10 to 100 numbers. A smaller effective filtering level, such as 0.1 or 0.2 times the standard deviation of the recorded

PPC values, should be selected for larger sample sizes (around 200). In order to compare complexity levels based on different effective filtering levels, or to track the complexity trend with increasing PPC records, normalization could be used to to make the results comparable. Finally, complexity level is not the only indicator for evaluating the performance of building construction systems; performance could also be associated with other indicators, such as scope to form an evaluation system for the thorough assessment of complex building construction systems.

### *7.5 Future Research*

This study has provided a platform and solid foundation for future research with the intention of exploring the complex nature of building construction systems and its measurements.

The literature review and theoretical analysis of complex building construction systems provided a comprehensive theoretical foundation for the future research of complex building construction systems. This is a new era for complex system and complex sciences in construction research and practice. The field of complexity allows the industry to look at itself from a different prism, a different point of view or paradigm. It requires practitioners to step out of the traditional way of doing things, borrow insights from other disciplines and industry, and experiment, learn and apply new methods in a continuous improvement atmosphere.

As this study was specific to finding the most efficient entropy algorithms out of the three commonly used ones, it sets the platform and stage for future studies to explore other complexity measurement algorithms, or combinations thereof, to be executed in similar fashion, thus testing new levels of influence.

As mentioned in the previous section, obtaining quantitative measurements of complexity levels, rather than just qualitative descriptions, provides the possibility and opportunity to explore and further analyze complexity in building construction. This provides us with the criteria to check different approaches in order to improve construction performance and reduce the complexity in building construction. It offers the metrics to compare different construction performances, different construction project delivery systems, or even different construction methods from the perspective of complexity and chaos theory. These could all be the subjects of future study.

REFERENCES


Abdelhamid, T. (2004). The self-destruction and renewal of lean construction theory: A prediction from Boyd's theory. In Proceedings of the 12th Annual Conference of the International Group for Lean Construction (IGLC12), Helsingore, Denmark.

Al-Angari, H. M., & Sahakian, A. V. (2007). Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. Biomedical Engineering, IEEE Transactions on, 54(10), 1900-1904.

Allen, P.M. (2008) The Importance of Complexity for the Research Agenda in the Building Environment. *Architectural Engineering and Design Management*, 4(1): 5-14.

Amaral, L.A.N., & Ottino, J.M. (2004) Complex Systems and Networks: Challenges and Opportunities for Chemical and Biological Engineers, *Chemical Engineering Science,* 59(8-9): 1653-1666.

Axelrod, R., & Cohen, M. (1999) Harnessing Complexity: Organizational Implications of a Scientific Frontier, New York: The Free Press.

Azimi, R. A., Lee, S., & AbouRizk, S. (2012). Applying basic control theory principles to project control: Case study of off-site construction shops. *Journal of Computing in Civil Engineering*, 26(6), 681-690.

Babloyantz, A. & Destexhe, A. (1988) Is the normal heart a periodic oscillator? *Biol. Cybern*, 58:203-211.

Baccarini, D. (1996). The Concept of Project Complexity – A Review. *International Journal of Project Management*, 14(4), 201-204.

Ballard, G. (1999). Improving work flow reliability. In Proceedings of the 7 th Annual Conference of the International Group for Lean Construction, 275-286.

Bandt, C., & Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical Review Letters*, 88(17), 174102.

Batten, D. F. (2001). Complex landscapes of spatial interaction. *The Annals of Regional Science*, 35(1), 81-111.

Begun, J.W., Zimmerman, B., & Dooley, K. (2003) Health care organizations as complex adaptive systems. In: Mick SM, Wyttenback M, eds. Advances in health care organization theory. San Francisco: Jossey-Bass.

Bennett, J., & Fine, B. (1980). Measurement of complexity in construction projects: Final report. Department of Construction Management, University of Reading.

Benbya, H., & McKelvey, B. (2006). Toward a complexity theory of information systems development. *Information Technology & People*, 19(1), 12-34.

Bertelsen, S. (2003) Complexity – Construction in A New Perspective. 11th Annual Conference in the International Group for Lean Construction. Blacksburg, VA.

Boltzmann, L. (1872) Weitere Studien über das Wärmegleich- gewicht unter Gasmolekülen. Sitzungsberichte der mathe- matisch-naturwissenschlaftlichen Classe der kaiserlichen Akademic der Wissenschaften Wien 66:275–370.

Boulding, K.E. (1956) General systems theory: the skeleton of science, *Management Science*, 3(3): 197-208.

Byrner, D. (1998) Complexity Theory and The Social Sciences: An Introduction. London: Routledge.

Cao, Y., Tung, W. W., Gao, J. B., Protopopescu, V. A., & Hively, L. M. (2004). Detecting dynamical changes in time series using the permutation entropy. *Physical Review-Series E*, 70(4; PART 2), 046217-046217.

Caracelli, V. J., & Greene, J. C. (1997). Crafting mixed‐method evaluation designs. New directions for evaluation, 1997(74), 19-32.

Chowdhury, S. (2001). The power of six sigma. Financial Times Prentice Hall.

Christodoulou, S. E. (2008). A bid‐unbalancing method for lowering a contractor's financial risk. *Construction Management and Economics*, 26(12), 1291-1302.

Christodoulou, S. E., Ellinas, G. N., & Aslani, P. (2009a). Disorder considerations in resource‐constrained scheduling. *Construction Management and Economics*, 27(3), 229-240.

Christodoulou, S. E., Ellinas, G., & Michaelidou-Kamenou, A. (2009b). Minimum moment method for resource leveling using entropy maximization. *Journal of Construction Engineering and Management*, 136(5), 518-527.

Cillier, P. (1998) Complexity and Postmodernism: Understanding Complex system. London: Routledge.

Cillier, P. (2002) Why We Cannot Know Complex Things Completely. *Emergence*, 4(1): 77–84.

Cilliers, P. (2005) Complexity, deconstruction and relativism. *Theory Cult Soc*, 22(5): 255–267.

Clausius, R. (1867). The mechanical theory of heat: with its applications to the steam-engine and to the physical properties of bodies. J. van Voorst.

Cohen, J., & Stewart, I. (2000). The collapse of chaos: Discovering simplicity in a complex world. Penguin UK.

Crawford, T. W., Messina, J. P., Manson, S. M., & O'Sullivan, D. (2005). Complexity science, complex systems, and land-use research. *Environment and Planning B: Planning and Design*, 32(6), 792-798.

Curlee, W., & Gordon, R. L. (2011). Complexity theory and project management. Wiley. com.

Daw, C. S., Finney, C. E. A., & Tracy, E. R. (2003). A review of symbolic analysis of experimental data. *Review of Scientific Instruments*, 74(2), 915-930.

Dekker, S., Bergstrom, J., Amer-Wahlin, I., & Cilliers, P. (2013) Complicated, Complex, and Compliant: Best Practice in Obstetrics. *Cognition Technology & Work*, 15(2): 189-195.

Del Cano, A., & de la Cruz, M. P. (2002). Integrated methodology for project risk management. *Journal of Construction Engineering and Management*, 128(6), 473-485.

De Meyer, A., Loch, C. H., & Pich, M. T. (2002). From variation to chaos. *MIT Sloan Management Review*, 43, 60-7.

Dewey, T. G. (1997). "Algorithmic complexity and thermodynamics of sequence-structure relationships in proteins." *Physical Review E,* 56(4), 4545.

Dixit, M. K., Fernández-Solís, J. L., Lavy, S., & Culp, C. H. (2010). Identification of parameters for embodied energy measurement: A literature review. *Energy and Buildings*, 42(8), 1238-1247.

Dubois, A., & Gadde, L. E. (2002). The construction industry as a loosely coupled system: implications for productivity and innovation. *Construction Management & Economics*, 20(7), 621-631.

Elnashaie, S. S., and Grace, J. R. (2007). "Complexity, bifurcation and chaos in natural and man-made lumped and distributed systems." *Chemical Engineering Science,* 62(13), 3295-3325.

Eshragh, A., Filar, J., & Nazar, A. (2011) A Projection-Adapted Cross Entropy (PACE) Method for Transmission Network Planning. *Energy System*, 2: 189-208.

Faulkner, B., & Russell, R. (1997). Chaos and complexity in tourism: In search of a new perspective. *Pacific Tourism Review*, 1(2), 93-102.

Feagin, J. R., Orum, A. M., & Sjoberg, G. (Eds.). (1991). A case for the case study. UNC Press Books.

Fernández-Solís, J. L. (2008). The systemic nature of the construction industry. *Architectural Engineering and Design Management*, 4(1), 31-46.

Fernández-Solís, J. L. (2013). Building construction: A deterministic non-periodic flow– A case study of chaos theories in tracking production flow. *Architectural Engineering and Design Management*, 9(1), 21-48.

Fernández-Solís, J.L., Rybkowski, Z. K., Xiao, C., Lü, X., Chae, L. S. (2014) General Contractor project of projects – a meta-project: Understanding the new paradigm and its implications through the lens of entropy. *Architectural Engineering and Design Management*, accepted for publication.

Fitsilis, P. (2009) Measuring the complexity of software projects. In Computer Science and Information Engineering, *2009 WRI World Congress*, (7), 644-648.

Geraldi, J., & Adlbrecht, G. (2007) On faith, fact, and interaction in projects. *Project Management Journal*, 38(1): 32- 43.

Geraldi, J. (2007) Patterns of complexity: the thermometer of complexity. *Project Perspectives*. XXIX: 4-9.

Gidado, K. I. (1996). Project complexity: the focal point of construction production planning. *Construction Management & Economics*, 14(3), 213-225.

González, V., Alarcón, L. F., & Mundaca, F. (2008). Investigating the relationship between planning reliability and project performance. *Production Planning and Control*, 19(5), 461-474.

Grabowski, F., & Strzalka, D. (2008) Simple, Complicated and Complex systems-The Brief Introduction. *2008 Conference on Human System Interactions*, 570-573.

Grassberger, P. (1988). Finite sample corrections to entropy and dimension estimates. *Physics Letters A*, 128(6), 369-373.

Gregory, A. H, (1999), What Is Lean Construction-1999, Proceedings IGLC-7 Seventh Conference of the International Group for Lean Construction, University of California Berkeley, California, USA, 1-10.

Harry, M., & Schroeder, R. (2000). Six sigma: the breakthrough management strategy revolutionizing the world's top corporations. Broadway Business, New York.

Heylighen, F., Cilliers, P., & Gershenson, C. (2007) Complexity and philosophy. In: Bogg J, Geyer R (eds) Complexity, science and society. Oxford: Radcliffe Publishing.

Hill, G. W. (1982) Group Versus Individual Performance Are N+1 Heads Betters Than One? *Psychological Bulletin*, 91(3): 517-539.

Homer-Dixon, T. (2002). The ingenuity gap: Facing the economic, environmental, and other challenges of an increasingly complex and unpredictable future. Random House Digital, Inc.

Kantz, H. (1994). A robust method to estimate the maximal Lyapunov exponent of a time series. Physics letters A, 185(1), 77-87.

Karimi-Hosseini, A., Haddad, O. B., & Mariño, M. A. (2011). Site selection of raingauges using entropy methodologies. *Proceedings of the ICE-Water Management*, 164(7), 321-333.

Kayastha, P. (2012). Application of fuzzy logic approach for landslide susceptibility mapping in Garuwa sub-basin, East Nepal. Frontiers of Earth Science, 6(4), 420-432.

Khandoker, A. H., Gubbi, J., & Palaniswami, M. (2008). Recognizing central and obstructive sleep apnea events from normal breathing events in ECG recordings. In Computers in Cardiology, 2008 (pp. 681-684). IEEE.

Kolmogorov, A. N. (1963). On tables of random numbers. Sankhyā: *The Indian Journal of Statistics, Series A*, 25(4), 369-376.

Koskela, L. (1992). Application of the new production philosophy to construction (No. 72). (Technical Report No. 72, Center for Integrated Facility Engineering, Department of Civil Engineering). Stanford, CA: Stanford university.

Koskela, L. J. (1999). Management of production in construction: a theoretical view. In Proceedings of the 7th Annual Conference of the International Group for Lean Construction. Berkeley, California, USA.

Koskela, L. J. (2000). An exploration towards a production theory and its application to construction. VTT Technical Research Centre of Finland.

Koskela, L. J., & Howell, G. (2002) The underlying theory of project management is obsolete. In Proceedings of the PMI Research Conference (pp. 293-302). PMI.

Kurths, J., Voss, A., Saparin, P., Witt, A., Kleiner, H. J., & Wessel, N. (1995). Quantitative analysis of heart rate variability. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1), 88-94.

Kwak, Y. H., & Anbari, F. T. (2006). Benefits, obstacles, and future of six sigma approach. Technovation, 26(5), 708-715.

Lean Construction Institution (2015). http://www.leanconstruction.org/training/glossary/#p.

Leedy, P. D., & Ormrod, J. E. (2005). Practical research. Planning and design,8.

Lewin, R., (1993) Complexity: Life at the Edge of Chaos. New York: Collier Books.

Li, J., & Ning, X. (2007). Classification of physiologic and synthetic heart rate variability series using base-scale entropy. *Physica A: Statistical Mechanics and its Applications*, 384(2), 423-428.

Linstone, H. A. (1999). Complexity science: implications for forecasting. *Technological Forecasting and Social Change*, 62(1), 79-90.

Lu, X., Clements-Croome, D., Viljanen, M. (2010) Integration of Chaos Theory and Mathematical Models in Building Simulation Part I; Literature Review. *Automation in Construction*, 19: 447-451.

Manson, S. M. (2001). Simplifying complexity: a review of complexity theory. *Geoforum*, 32(3), 405-414.

Mao, P., Cheng, H., & Wang, X.D. (2009) Study on Organization Structure of Programme Management Based on Information Entropy. *International Symposium on Advancement of Construction Management and Real Estate*, 1-6: 2439-2444.

McKercher, B. (1999) A Chaos Approach to Tourism. *Touriam Management*, 20(4): 425-434.

Myat, A., Thu, K., et al, Ng, K.C. (2012) Entropy Generation Minimization: A Practical Approach for Performance Evaluation of Temperature Cascaded Co-generation Plants. *Energy*, 46: 493-521.

Nicolis, G., and Prigogine, I. (1989). "Exploring complexity." *New York*, .

Oseledets, V. I. (1968) A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. Trans. Moscow Math. Soc, 19: 197-231.

O'Sullivan, D., Manson, S. M., Messina, J. P., & Crawford, T. W. (2006). Space, place, and complexity science. *Environment and Planning A*, 38(4), 611-617.

Page, S. E. (2008) Uncertainty, difficulty, and complexity, *Journal of Theoretical Politics*, 20(2), 115-149.

Pande, P., Neuman, R., & Cavanagh, R. R. (2000). The six sigma way: How GE, Motorola, and other top companies are honing their performance. McGraw Hill Professional.

Pande, P. S., Holpp, L., & Pande, P. (2002). What is six sigma? (Vol. 1). New York, NY: McGraw-Hill.

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of the Association of American Geographers*, 93(2), 314-337.

Pheng, L. S., & Hui, M. S. (2004). Implementing and applying Six Sigma in construction. Journal of construction engineering and management, 130(4), 482-489.

Pincus, S.M. (1991) Approximate Entropy as a Measure of System Complexity. *Proceeding of National Academic of Science*, 88: 2297-2301.

Radosavljević, M., & Horner, R. M. W. (2002). The evidence of complex variability in construction labor productivity. *Construction Management & Economics*, 20(1), 3-12.

Remington, K., Zolin, R., & Turner, R. (2009) A model of project complexity: distinguishing dimensions of complexity from severity. International Research Network of Project Management Conference. Berlin.

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.

Rosenhead, J. (1998). Complexity theory and management practice. Department of Operational Research, London School of Economics and Political Science.

Rounds, J. L., & Segner, R. O. (2010). Construction Supervision. Wiley. com.

Rubinstein, R.Y., (1997) Optimization of computer simulation models with rare events. Eur. J. Oper. Res, 99: 89–112.

Rubinstein, R. Y., & Kroese, D. P. (2004). The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning. Springer.

Rudolf, C. (1850) On the Motive Power of Heat, and on the Laws which can be deduced from it for the Theory of Heat. Poggendorff's Annalen der Physick, LXXIX (Dover Reprint). ISBN 0-486-59065-8.

Saparin, P., Witt, A., Kurths, J., & Anishchenko, V. (1994) The Renormalized Entropy - An Appropriate Complexity Measure? *Chaos, Solutions & Fractals*, 4(10): 1907-1916.

Serres, M., Harari, J. V., and Bell, D. F. (1982). *Hermes: Literature, science, philosophy.* Johns Hopkins University Press Baltimore, .

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 22(7), 379-423.

Sinai, Ya. G. (1959) Dokl. Akad. Nauk SSSR 124, 768-771.

Snee, R. D. (2000). Impact of Six Sigma on quality engineering. Quality Engineering, 12(3), 9-14.

Spencer, H. (1887). The genesis of science. Humboldt Publishing Company.

Stevenson, H., & Harmeling, S. (1990) Entrepreneurial Management's Need for A More "Chaotic" Theory. *Journal of Business Venturing*, 5: 1-14.

Stwart, I., (2002) Does God Play Dice?-The new Mathematics of Chaos (2nd ed). Malden, MA: Blackwell Publishing Ltd.

Tang, L. C. M., Leung, A. Y. T., & Wong, C. W. Y. (2009). Entropic risk analysis by a high level decision support system for construction SMEs. *Journal of Computing in Civil Engineering*, 24(1), 81-94.

Tang, P., Mukherjee, A., & Onder, N. (2010). Studying dynamic decision-making in construction management using adaptive interactive simulations. In Construction Research Congress (pp. 379-38).

Tatikonda, M.V., & Rosenthal, S.R. (2000) Technology novelty, project complexity, and product development project execution success: a deeper look at task uncertainty in product innovation. *IEEE Transactions on Engineering Management*, 47(1): 74-87.

Teddlie, C., & Tashakkori, A. (2003). Major issues and controveries inthe use of mixed methods in the social and behvioral sciences. Handbook of mixed methods in social & behavioral research, 3-50.

Thomas, J., & Mengel, T. (2008) Preparing Project Managers to Deal With Complexity - Advanced Project Management Education. *International Journal of Project Management*, 26: 304-315.

Thompson, J.M.T., & Gray, P. (1990) Chaos and Dynamic complexity. London: Royal Society.

Turner, J. R., & Cochrane, R. A. (1993) Goals-and-methods matrix: coping with projects with ill defined goals and/or methods of achieving them. *International Journal of Project Management*, 11 (2): 93-102.

Vidal, L. A., & Marle, F. (2008). Understanding project complexity: implications on project management. *Kybernetes*, 37(8), 1094-1110.

Vidal, L.A., Marle, F., & Bocquet, J.C. (2011) Measuring Project Complexity Using the Analytic Hierarchy Process. *International Journal of Project Management*, 29: 718-727.

Waldrop, M.M. (1992) Complexity: the emerging science at the edge of order and chaos, London: Viking.

Wiener, N. (1948). Cybernetics (p. 112). Paris: Hermann.

Williams, T. M. (1999). The need for new paradigms for complex projects. *International Journal of Project Management*, 17(5), 269-273.

Williams, T.M. (2005) Assessing and moving on from the dominant project management discourse in the light of project overruns. *IEEE Transactions on Engineering Management*, 52(4): 497-508.

Winch, G. (1989). The construction firm and the construction project: a transaction cost approach. *Construction Management and Economics*, 7(4), 331-345.

Wolf, A., Swift, J. B., Swinney, H. L., & Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. Physica D: Nonlinear Phenomena, 16(3), 285-317.

Xu, X., Zhang, J., & Small, M. (2008). Superfamily phenomena and motifs of networks induced from time series. *Proceedings of the National Academy of Sciences*, 105(50), 19601-19605.

Yugue, R.T., & Maximiano, A.C.A. (2012) Contribution to the Research of Project Complexity and Management Processes. Management of Innovation and Technology (ICMIT), *2012 IEEE International Conference*, 668-673.

Zbilut, J. P., Mayer-Kress, G., & Geist, K. (1988). Dimensional analysis of heart rate variability in heart transplant recipients. *Mathematical Biosciences*, 90(1), 49-70.

Zeng, X., Eykholt, R., & Pielke, R. A. (1991). Estimating the Lyapunov-exponent spectrum from short time series of low precision. Physical Review Letters, 66(25), 3229.

Zhang, S., Zhao, Y., Jia, J, & Shangguan, H. (2010) Research on the Chaos Recognition Method Based on Differential Entropy. *Chinese Physcial B*, 19(6): 060514.

Zhang, T. L., Ding, Y. S., & Chou, K. C. (2008). Prediction protein structural classes with pseudo-amino acid composition: approximate entropy and hydrophobicity pattern. *Journal of theoretical biology*, 250(1), 186-193.

Zheng, Y.L., & Li, Q.M. (2009) Study on Application of Entropy Theory in Construction Project Communication. International *Symposium on Advancement of Construction Management and Real Estate*, 1-6:2021-2028.

Zimmerman, B., Lindberg, C., & Plsek, P. (1998) Edgeware: insights from complexity science for health care leaders. Irving, TX: Veterans Health Affairs.

Zimmerman, B., Lindberg, C., & Plsek, P. (2000). A complexity science primer: What is complexity science and why should I learn about it. Edgeware: Insights from Complexity Science for Health Care Leaders.

APPENDIX A

## MATLAB code for Approximate Entropy (ApEn) Algorithm

```
function apen = ApEn( dim, r, data, tau )
% ApEn
% dim: embedded dimension
%  r: tolerance (typically 0.2 * std)
%  data: time-series data
%---------------------------------------------------------------

N = length(data);
result = zeros(1,2);

for j = 1:2
   m = dim+j-1;
   phi = zeros(1,N-m+1);
   dataMat = zeros(m,N-m+1);

   % setting up data matrix
   for i = 1:m
      dataMat(i,:) = data(i:N-m+i);
```

```
    end

    % counting similar patterns using distance calculation
    for i = 1:N-m+1
        tempMat = abs(dataMat - repmat(dataMat(:,i),1,N-m+1));
        boolMat = any( (tempMat > r),1);
        phi(i) = sum(~boolMat)/(N-m+1);
    end

    % summing over the counts
    result(j) = sum(log(phi))/(N-m+1);
end

apen = result(1)-result(2);
end
```

APPENDIX B

## MATLAB code for Sample Entropy (SampEn) Algorithm

```
function sampen = SampEn( dim, r, data )


%   dim: embedded dimension

%   r: tolerance (typically 0.2 * std)

%   data: time-series data

%-----------------------------------------------------------------


N = length(data);

correl = zeros(1,2);

dataMat = zeros(dim+1,N-dim);

for i = 1:dim+1

    dataMat(i,:) = data(i:N-dim+i-1);

end


for m = dim:dim+1

    count = zeros(1,N-dim);

    tempMat = dataMat(1:m,:);
```

```
    for i = 1:N-m
        % calculate Chebyshev distance, excluding self-matching case
        dist = max(abs(tempMat(:,i+1:N-dim) - repmat(tempMat(:,i),1,N-dim-i)));

        % calculate Heaviside function of the distance
        % User can change it to any other function
        % for modified sample entropy (mSampEn) calculation
        D = (dist < r);

        count(i) = sum(D)/(N-dim);
    end

    correl(m-dim+1) = sum(count)/(N-dim);
end

sampen = log(correl(1)/correl(2));
end
```

APPENDIX C

**MATLAB code for Permutation Entropy (PEn):**

```
function [pe hist] = PEn(y,m)


%  Input:   y: time series;
%              m: order of permuation entropy
% Output:
%           pe:    permuation entropy
%           hist:  the histogram for the order distribution
%-----------------------------------------------------------------------


ly = length(y);
permlist = perms(1:m);
c(1:length(permlist))=0;


 for j=1:ly-(m-1)
    [a,iv]=sort(y(j:1:j+(m-1)));
    for jj=1:length(permlist)
        if (abs(permlist(jj,:)'-iv))==0
            c(jj) = c(jj) + 1 ;
```

```
        end
    end
end

hist = c;


c=c(find(c~=0));
p = c/sum(c);
pe = -sum(p .* log(p));
```

APPENDIX D

## MATLAB Code for Maximal Lyapunov Exponent

```
function lam = lyapunov(y,tau)

% calculate the largest positive Lyapunov exponent from time series data

if tau==0;

%____________________Determination of Embeding Lag: tau______________________


% A: Autocorrelation


y=y(:);

[nyr,nyc]=size(y);


[ACF,Lags,Bounds] = autocorr(y(:,1),10,[],[]);

ACF=ACF(2:end);

for l=1:10

   if abs(ACF(l))<=exp(-1),tau=l-1; break,end

end


if tau==0
```

```
% B:  Minimum Mutual Information
pnts=100;
for im=0:10
   z=lagmatrix(y,im);
d=2;
n=length(z(im+1:end));

endp1=ceil(pnts/10);
endp2=ceil(pnts/10);

minz=min(z(im+1:end));maxz=max(z(im+1:end));grz=(maxz-minz)/(pnts-endp1);
miny=min(y(im+1:end));maxy=max(y(im+1:end));gry=(maxy-miny)/(pnts-endp1);

h1z=(4/(3*n))^(1/5)*std(z(im+1:end));
h1y=(4/(3*n))^(1/5)*std(y(im+1:end));

for k=1:pnts
zi(k,1)=minz+grz*(k-endp2);
yi(k,1)=miny+gry*(k-endp2);

fz(k,1)=(1/((2*pi)^0.5*n*h1z))*sum(exp(-((zi(k,1)-...
   z(im+1:end)).^2)/(2*h1z^2)));
```

```
fy(k,1)=(1/((2*pi)^0.5*n*h1y))*sum(exp(-((yi(k,1)-...
   y(im+1:end)).^2)/(2*h1y^2)));


pz(k,1)=(1/((2*pi)^0.5*n*h1z))*sum(exp(-((zi(k,1)-...
   z(im+1:end)).^2)/(2*h1z^2)))*grz;
py(k,1)=(1/((2*pi)^0.5*n*h1y))*sum(exp(-((yi(k,1)-...
   y(im+1:end)).^2)/(2*h1y^2)))*gry;


end


[gz gy]=meshgrid(zi,yi);
sigma=((n*var(z(im+1:end))+n*var(y(im+1:end)))/(n+n))^0.5;
h=sigma*(4/(d+2))^(1/(d+4))*(n^(-1/(d+4)));


for i=1:pnts
   for j=1:pnts


      fzy(i,j)=(1/(2*pi*n*h^2))*sum(exp(-((gz(i,j)-z(im+1:end)).^2+...
         (gy(i,j)-y(im+1:end)).^2)/(2*h^2)));
      pzy(i,j)=(1/(2*pi*n*h^2))*sum(exp(-((gz(i,j)-z(im+1:end)).^2+...
         (gy(i,j)-y(im+1:end)).^2)/(2*h^2)))*grz*gry;
      I1zy(i,j)= pzy(i,j)*log(pzy(i,j)/(pz(i)*py(j)));
```

```
    end
end

Hz=-(pz'*log(pz));
Hy=-(py'*log(py));

MIzy=(sum(sum(I1zy)));
RMIzy1(im+1,1)=2*MIzy/(Hz+Hy);
RMIzy2(im+1,1)=MIzy/(Hz*Hy)^0.5;
RMIzy3(im+1,1)=MIzy/min(Hz,Hy);
end

MIInd=find(RMIzy1(2:end)<exp(-1)*RMIzy1(1,1));
if MIInd(1,1)>1
tauMI=MIInd(1,1)-1;
else
   tauMI=1;
end

tau=tauMI;

end
```

```
end

[ndata nvars]=size(y);

N2 = floor(ndata/2);
N4 = floor(ndata/4);
TOL = 1.0e-6;

exponent = zeros(N4+1,1);

for i=N4:N2  % second quartile of data should be sufficiently evolved
  dist = norm(y(i+1,:)-y(i,:));
  indx = i+1;
  for j=1:ndata-5
     if (i ~= j) && norm(y(i,:)-y(j,:))<dist
        dist = norm(y(i,:)-y(j,:));
        indx = j; % closest point!
     end
  end
  expn = 0.0; % estimate local rate of expansion (i.e. largest eigenvalue)
  for k=1:5
     if norm(y(i+k,:)-y(indx+k,:))>TOL && norm(y(i,:)-y(indx,:))>TOL
```

```
        expn = expn + (log(norm(y(i+k,:)-y(indx+k,:)))-log(norm(y(i,:)-y(indx,:))))/k;
      end
  end
  exponent(i-N4+1)=expn/5;
end

% plot(exponent);  % plot the estimates for each initial point (fairly noisy)

summ=0;  % now, calculate the overal average over N4 data points ...
for i=1:N4+1
   summ = summ+exponent(i);
end

lam=summ/((N4+1)*tau);  % return the average value
% if lam > 0, then system is chaotic
```