BAYESIAN ASSESSMENT OF REGIONAL OIL AND GAS PRODUCTION

A Dissertation

by

YASSER SOLTANPOUR

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,	Zenon Medina-Cetina
Committee Members,	Walter Ayers
	Thomas A. Blasingame
	Ivan Damnjanovic
	Eduardo Gildin
Head of Department,	Robin Autenrieth

December 2017

Major Subject: Civil Engineering

Copyright 2017 Yasser Soltanpour

ABSTRACT

Several empirical or analytical/semi-analytical simulation models have been developed to assess the Estimated Ultimate Recovery (EUR) of an oil or gas formation for a short or long term of production. Furthermore, providing the EUR of a set of regional wells, it often becomes essential to perform a spatial analysis to develop the overall perception of possible depletion rate across the play. However, the lack of knowledge regarding the likely statistical structure of simulation models' parameters coupled with the unknown influence of correlation amidst wells' locations, makes it pertinent to apply a mechanism to quantify the uncertainty associated with the analysis. Therefore, in this study, researchers initially exerted the principles of the Bayesian paradigm together with the Markov Chain Monte Carlo (MCMC) theory to capture the posterior of the simulation model random field. Also, a vector of randomly drawn samples from the retrieved posterior allows delineation of the expected model realizations for a course of progressive time.

Despite the fact that MCMC incorporating the acceptance-rejection criterion of the Metropolis-Hastings (MH) algorithm eventually converges to the true mean of the random process, it appears that the general trend of sampling often suffers from being computationally inefficient. Accordingly, to address the aforementioned issue, a novel sophisticated framework which is called "Parallel Scaled Adaptive Metropolis-Hastings" is developed. PSAMH constructs several synchronous chains to adapt the step size of MH proposal distribution and hence optimize the acceptance rate.

Moreover, in this study, three major EUR evaluation techniques are employed. The Power Law Exponential Decline (PLED) and Modified Hyperbolic Decline (MHD) functions, along with a semi-analytical method, serve to project the well production performance over the varying time. Additionally, the depletion logs given from the Eagle Ford Shale and Barnett Shale deliver the required observation data. Besides, the Ordinary Kriging and Inverse Distance Weight are two key techniques that are applied to approximate the spatial behavior of the formation.

In addition, researchers elaborated a sequential Bayesian updating mechanism to take the updating evidence into the prior's computation for various time intervals. Also, a Bayesian-spatial algorithm is used to feature the spatial characteristics of unexplored locations hypothesizing the fact that the only given information comprises the production observed data and corresponding coordinate for each individual well.

It is implied that exerting the Bayesian approach permits quantifying the inherent uncertainty in the model analysis. Furthermore, it is concluded that the sequential Bayesian updating mechanism is able to noticeably increase the performance and efficiency of the process by precisely constructing an appropriate prior distribution. Also, it is connoted that, given merely the observation data, associated coordinates and EUR evaluation models, it becomes possible to estimate the statistics of model variables and the production behavior for different courses of time at desired locations. Last but not least, attaining the Bayesian-spatial production forecasting for varying depletion times, it becomes plausible to generate the daily basis and cumulative production dynamic maps.

DEDICATION

This dissertation is firstly dedicated to my wife, Shima, for all her unconditional love, support, motivation and patience throughout not only my PhD education but also in our entire marriage life. Secondly, it is dedicated to my mother, Shahla, my father, Ahmad, and my siblings, Fareed and Leila for their support and love. It is an irrefutable fact that without my wife and parents, I would not be able to make this moment happen. Finally, it is dedicated to the life, itself, and the faith which has brought me this moment.

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Zenon Medina-Cetina, for his mentoring, teaching and encouragements and my committee members, Dr. Ayers, Dr. Blasingame, Dr. Damnjanovic, and Dr. Gildin, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience.

I also would like to extent my sincere to Dr. Gorgonio Fuentes-Cruz and Ms. Nefeli Moridis for their help about the problems relevant to the reservoir engineering.

Additionally, I would like to appreciate the supports provided by Dr. Alberto Muñoz and Mr. Victor Gutierrez from Grupo Plenum México.

Finally, thanks to my mother and father for their encouragement and to my wife for her great sacrifice, patience and love.

CONTERIBUTORS AND FUNDING SOURCES

This work was supervised by a dissertation committee consisting of committee chair, Professor Zenon Medina-Cetina and Professor Ivan Damnjanovic of the Department of Civil Engineering and Professor(s) Walter Ayers, Thomas A. Blasingame and Eduardo Gildin of the Department of Petroleum Engineering.

The data analyzed from the Eagle Ford Shale was provided by Professor Thomas A. Blasingmae and Ms. Nefeli Moridis and belongs to Murphy Oil Corporation. The data analyzed from the Barnett Shale was provided by Dr. Gorgonio Fuentes, and was taken from the Drillinginfo website.

The research was funded by Grupo Plenum México and in particular by Dr. Alberto Muñoz and Mr. Victor Gutierrez. All work conducted for the dissertation was completed by the student independently.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

The Bayesian paradigm is defined by the influence of the combination of the tradeoff between observations and model predictions (Likelihood) and the impact of the prior knowledge about model parameters (Prior) (Bayes and Price 1763; Vidakovic 1962; Gelman et al. 2013). The implementation of the Bayesian paradigm is also known as a probabilistic inversion or a probabilistic solution to an inverse problem. The resulting distribution from a Bayesian analysis is called the Posterior, from which inferences can be generated to assess the model parameters, and consequently the model predictions. A Bayesian formulation accounts by a) the number of random parameters, b) the accuracy and complexity of predictive models and c) the amount of observed data from the process of interest. In addition, it populates the correlation structure among the model parameters, and provides statistical inferences that increases the apprehension of the model behavior by accounting for the uncertainty inherent to the available evidence (a, b and c from above) (Gelman et al. 2013).

Additionally, the Markov Chain Monte Carlo technique (MCMC) (Berg and Billoire 2008; Gilks and Richardson 1996; W. Al-mudhafar 2015) coupled with Metropolis-Hastings (MH), Random Walk Metropolis (RWM) algorithm (Metropolis et al. 1953; Hastings 1970) and Gibbs sampler (Gelman et al. 2013) as subordinates of the Bayesian paradigm, have recently become common practices dealing with complicated probabilistic problems. MCMC integrates thousands of randomly generated samples to eventually

converge to the mean of random process, utilizing the acceptance-rejection criterion of the MH algorithm. Cipra (2000), in addition, listed the Metropolis-Hastings algorithm among the top 10^{th} remarkable algorithms in 20^{th} century. MH algorithm draws random samples through an auxiliary distribution, which is denoted as “Proposal distribution”, “Kernel MCMC” or “Transition Kernel” (Sims 1998; Haario, Saksman, and Tamminen 2001; Gareth O. Roberts and Rosenthal 2002; Xifara et al. 2014). However, implementing MCMC applying the MH algorithm requires drawing random samples with a considerable chance of rejection. Accordingly, an undetermined and long convergence time, reasonably, initiates a critical drawback making MCMC computationally inefficient. That is for, inevitable argues have been sparked among scholars concerning the efficiency of MH or RWM algorithms. So far, a number of attempts have been conducted to overcome the convergence pitfall. However, a robust framework that consistently optimizes the machine energy and running time without significantly increasing the complexity of the problem, yet to be addressed. Typically, recent developments deal with the MCMC convergence efficiency regulating two constitutive approaches. Several methods, in one side, expedite the convergence with optimizing or tuning the proposal distribution (Sims 1998; Haario, Saksman, and Tamminen 2001; Gareth O. Roberts and Rosenthal 2002; Xifara et al. 2014), whilst in the other side, some other frameworks improve capturing the target distribution by exploiting an ensemble of likely distributions (Kendall, Liang, and Wang 2005b; Wang and Swendsen 2004; Liu, Liang, and Wong 2000; Earl and Deem 2005). Nevertheless, it is investigated that a combination of both strategies can be a solution to the MH drawback.

Adaptive Markov Chain Monte Carlo (Adaptive-MCMC) (Rosenthal 2011; Graves 2011; Andrieu and Thoms 2008; Sims 1998) and Adaptive Metropolis Hastings are two methods that allow us to tune the step size of the MH proposal distribution and hence expedite the convergence time. The acceptance-rejection criterion of random samples serves as the core controller in Adaptive-MCMC which is frequently invoked as the acceptance rate. Nonetheless, applying the Adaptive-MCMC method, one requires to utilize several trials and errors to manually adjust the step size and optimize the acceptance rate. Furthermore, it is evident that an acceptance rate amidst 0.44 and 0.234 optimally maps the RWM algorithm when the size of random space fluctuates between one to infinity. Also, note that, selection of an appropriate proposal distribution appears as another significant property of adaptive methods which should be properly addressed. Ideally, the best proposal distribution is the target distribution (Posterior) itself, nevertheless, in MCMC, presumably, it is either impossible or extremely difficult to sample directly from the target distribution.

In accordance to the aforementioned Adaptive-MCMC difficulties, in this study, a novel framework is constructed that by applying several synchronous chains aims to adapt the step size of the proposal distribution to eventually optimize the acceptance rate. Note that, although, the parallelizing concept has previously employed in the Parallel Tempering (PT) (Wang and Swendsen 2004; Earl and Deem 2005) and Multiple Tries Metropolis (MTM) (Liu, Liang, and Wong 2000) , we merged the similar notion into the genuine structure of Adaptive-MCMC to develop a robust and sophisticated method which is called the “Parallel Scaled Adaptive Metropolis-Hastings” (PSAMH) framework. It

should be noticed that, any method that exerts the parallel chain technique as the core mechanism, performs efficiently when the computational model which the samples are drawn from that is able to be iterated substantially fast.

Several well-known models allow practitioners to evaluate the Estimated Ultimate Recovery (EUR) of conventional or unconventional oil or gas reserves. Generally, three groups of EUR acquisition approaches are developed (Arps 1944; Ilk et al. 2008), which are categorized into the history matching, empirical decline curve (Arps) and analytical/semi analytical decline curve models. Provided that, three key EUR simulation methods are proposed in this work comprising two empirical methods, the Modified Hyperbolic Decline (MHD) and Power Law Exponential Decline (PLED) curve (Seshadri and Mattar 2010; Ilk et al. 2008) and the "Hydraulically Fractured wells based on the Non-uniform Induced Properties" (HFNIP) model (Fuentes-Cruz, Gildin, and Valkó 2014) which incorporates the hydro-chemo-mechanical characteristics of a play into calculation . The production flow rate units for empirical models is defined as Barrels of Oil Equivalent per Day (BOED), whilst the semi-analytical framework drives the computation in Thousand Standard Square Feet per Day (Mscf/D). Hereafter and according to the Bayesian literature, the EUR simulation models are referred as "forward model". Additionally, to be able to assign appropriate values to the forward model parameters, often several trials and errors require to be driven incorporating subjective reasoning. To diminish the subjectivity of assigning process, this study constitutes a methodology to assess the forward model parameters by the use of the Bayesian paradigm, which allows estimating the first and second order statistics of random field. Moreover, the posterior

vector retrieved from the probabilistic analysis, makes it possible to delineate the likely model realizations for various production time. In addition, recently, Gong et al. (2011),(2014); Cheng et al. (2010) and Moridis et al. (2017) exerting the model realizations have directly quantified the associated uncertainty within the oil or gas well production.

A key characteristic of the Bayesian paradigm appears in the updating capability of the random field in the light of the new available evidence (i.e. a) experimental observations b) model complexity c) expert beliefs). Therefore, we suggested a methodology to investigate the influence of adding new evidence on the performance of Bayesian analysis, and in particular the uncertainty quantification. That is for, an algorithm is constructed to readily generate a hybrid prior to not only take the advantage of retrieved posteriors from previous MCMC computations but also be able to freely explore the entire random space by switching to a more broaden prior when the employed posterior imposed unnecessary constraints.

Also, in the field of oil and gas, it is intuitively known that, decision-makers frequently encounter numerous challenges that require an abstract and quick treatment regarding the next investment step. To deal with similar challenges, providing a reliable production data, expert's beliefs and precise prediction models appear as practical functions to support their decisions. To elucidate the issue, a relevant example turns to be the possible location of the next drilling within the in-potential oil and gas reserves (Sidler, Prof, and Holliger 2003). In such cases, it is a common practice to employ the current available data to determine the possible potential of formation at unexplored coordinates, exerting the previously provided production in proximity of regional wells. Therefore,

taking advantage of applying the previous probabilistic inversion statistics of interested region is one of the economical methods which not only preserves the exploration budget of a project but also substantially enlightens the direction of decision. Despite the significant developments have recently been occurred in the technology, yet, the complexity and confusion associated with the accuracy of provided data need to be addressed. A practical method that is able to increase the precision of the analysis over the acquired data is the spatial geo-statistics analysis of model parameters and flow rate prediction (Fotheringham, Charlton, and Brunsdon 1998; Weber and Englund 1992). Spatial mapping or spatial statistics, studies the topological, geometric or geographic properties of target entities via formal techniques (Lehmann, Overton, and Leathwick 2002; Mitas and Mitasova 1999). In other words, the spatial analysis provides a reliable tool to develop the current knowledge about the unknown property at unexplored coordinates through interpolation amidst known quantities of specific parameter(s) (Caruso and Quarta 1998). Therefore, obtaining Bayesian statistics (Banerjee and Fuentes 2012) for a short and long term production together with the corresponding well coordinates allow us to derive the spatial characteristics of the target entity and thus enables us to delineate a dynamic map over the desired region. This goal has become achievable exploiting the geo-statistics data analysis and in particular the Kriging approach(Brown and Falade 2003; Bohling 2005b). Kriging which is identified as one of the most widely applied spatial analysis techniques, splits the surface into segments aims to govern the probabilistic analysis over entire area. Besides, the Inverse Distance Weight

is another reliable spatial method that has additionally been utilized in this study (Yasrebi et al. 2009).

The current context outlines several sections incorporating under-publication progress papers. Initially, an introduction into the PSAMH framework, implementation techniques, augmentation features and two numerical examples are elaborated in sections 2 and 3. The application of PSAMH on the empirical simulation models, MHD and PLED, exerting the Eagle Ford Shale data and a comparison amid the mentioned functions are delivered in section 4. In section 5, we employed the semi-analytical model in association with the Barnett Shale data to quantify the uncertainty inherent in simulation computations. The updating Bayesian methodology along with the PSAMH algorithm is comprehensively delineated in section 6. Then using the provided Bayesian inferences, a thoroughly comparison derived between four wells of the Eagle Ford Shale applying the MHD empirical model. Section 7 encompasses both Bayesian paradigm coupled with the PSAMH and the spatial analysis technique exerting the observed data attained from 43 wells of the Eagle Ford Shale development. Some concise remarks of entire previous sections are recapitulated in section 8. The results of Bayesian and spatial analysis computed in section 7, are tabulated in Appendices A and B, respectively.

2. PRACTICAL BAYESIAN FRAMEWORK TO PARALLELIZE ADAPTIVE MARKOV CHAIN MONTE CARLO: MECHANISM AND APPLICATION

2.1. Overview

The full integration of posterior distribution applying Markov Chain Monte Carlo (MCMC) and Metropolis-Hastings (MH) algorithm in Bayesian analysis requires drawing thousands of samples prior the stationary, which often turns to be computationally inefficient. So far, several methods have been constructed to deal with this inefficiency; however, they often increase the order of complexity of MCMC implementation and also entail acquisition a comprehensive knowledge in Bayesian techniques. Therefore, this study aims to develop a practical framework to optimize the convergence of MCMC needed to formulate the Bayesian inference. Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) exerting the adaptive MCMC methodology, randomly draws synchronous chains and automatically tunes the step size of proposal distribution to attain the optimum MH acceptance rate, whilst eventually ensures capturing all posterior modes in a more computationally efficient approach. Additionally, a synthetic experiment is employed to delineate the implementation sequence of PSAMH framework. The results imply that PSAMH by efficiently exploring the entire random field is capable to precisely capture all posterior modes in a more reasonable computational time.

2.2. Introduction

The Bayes theorem has recently become a common tool to develop the statistical inferences with the engineering applications. Whenever, a practitioner takes the prior knowledge of an event into the calculation, the Bayes theorem is unconsciously exerted. One simple example of the application of Bayesian theorem is “Bayesian Spam filter” in the computer science which provides a mechanism to detect a spam email according to the prior pattern of received emails in the daily basis routine. Another application appears in the risk assessment whereas the term of “vulnerability” and “hazard” are in fact conditional probability quantities.

In this study, a practical adaptive Markov Chain Monte Carlo (AMCMC) framework (Sims 1998; Andrieu and Thoms 2008; Atchadé et al. 2009; Rosenthal 2011; Graves 2011) is proposed aimed to automatically tune the step size of the proposal distribution in the Metropolis-Hastings (MH) (Metropolis et al. 1953; Hastings 1970) algorithm. The Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) is a robust algorithm that draws several concurrent chains to optimize the acceptance rate, whereas the acceptance rate indicates the ratio of the accepted to the total number of generated samples in the MH algorithm.

The Bayesian inference assesses the current state of belief by relying on the conditional probability of the physical model and observation (Hoff 2009; Gelman et al. 2013). That is for, the Bayesian paradigm incorporates the tradeoff between the observed data and physical model results (known as the Likelihood) and the prior knowledge of the model parameters as the evidence to draw the posterior distribution. The product of likelihood and prior distributions is often addressed as the target distribution and the

physical model, according to the Bayesian literature, is invoked as “forward model”. However, the posterior distribution within the Bayesian paradigm is often difficult to sample directly, therefore using the method of Markov Chain Monte Carlo (MCMC) permits generating random samples from the target distribution to capture the mean of random process.

MCMC (Gilks and Richardson 1996; Geyer 2002; Faming, Chuanhai, and Raymond 2010; Berg and Billoire 2008) coupled with the Metropolis-Hastings (MH) algorithm are subordinates of the Bayesian paradigm which have recently found their applications in the probabilistic analysis due to their capability in dealing with complicated posteriors. MCMC integrates thousands of drawn random samples to converge to the expected value of the random process by essentially accounting merely on a generated sample prior to the current state. It is, therefore, assumed that the generated random samples forget the influence of starting point when the number of samples significantly increases, which is also known as the Markovian state of MCMC. Additionally, MH, on the other hand, provides the required criterion to accept or reject the generated samples when it is impossible to directly sample from the target distribution. MH samples an auxiliary distribution, which is called proposal or kernel distribution, and computes the probability of proceeding the chain. Random Walk Metropolis (RWM), moreover, drives the MH criterion when the proposal distribution is symmetrical. However, generating Markov Chains through MH requires drawing random samples with a noticeable chance of rejection. Furthermore, despite an irrefutable convergence to the true posterior via MCMC, an inherent dependency of generated samples and an unknown convergence time

are critical drawbacks which make scholars be reluctant in applying this method. Researchers, so far, have formulated a number of approaches to diminish the aforementioned difficulties.

In general, most alternative methods sample the posterior space by either tuning the step size of the proposal distribution or altering the target distribution (Kendall, Liang, and Wang 2005a). Adaptive Markov Chain Monte Carlo (AMCMC) and Adaptive Metropolis (AM) (Haario, Saksman, and Tamminen 1999; Haario, Saksman, and Tamminen 2001) are two well-known methods that contribute the correlation structure amidst random variables by imposing the empirical covariance into random sampling when they become available. These techniques are required to deterministically tune the step size of the proposal distribution to optimize the acceptance rate. Gelman et al. (1996) and Roberts et al. (1997) proved that an acceptance rate of 0.234 optimally maps the RWM algorithm when the size of random field tends to infinity and also gets 0.44 when the size of parameter space becomes one. In addition, they mainly assumed that the experiment acceptance rate is only a function of parameter space size. Furthermore, the dependency on the previously generated samples, has become a considerable drawback inherent in Adaptive methods that eventually invalidates the fundamental hypothesis of the MCMC Markovian state in Adaptive samplers. Also, when adaptive methods are employed, it appears necessary to investigate the state of ergodicity (reversibility of MCMC) of sampler case by case (Craiu et al. 2014). Andrieu and Thoms (2008) and Roberts and Rosenthal (2009) immersed deeply into the application and implementation of the adaption theory in the MH algorithm and recommended several remedies to diminish adaption method's

pitfalls. Another augmentation on adaptive technique introduced by Craiu et al. (2009) when the notion of AMCMC merged with the parallel sampling notion by running several parallel chains and recalling them after a predefined batch size iterations.

Provided that, this work by advocating the aforementioned techniques elaborates the PSAMH framework that automatically tunes the step size of the proposal distribution and optimizes the acceptance rate empowered by parallelizing independent synchronous chains. The proposal distribution is modified subsequently after a fix number of iterations (batch sizes) such that it enables the sampler to freely explore the entire parameter space in order to capture all possible posterior modes. The key technique used in the sampling, initiates from developing ensembles of random coefficient of variation (CoV) in the parallel chains. The coefficient of variation (CoV) (Everitt 2006; Forkman and Verrill 2008; Forkman 2009) exhibits the state of dispersion of data with respect to the mean of samples. Since the step size is, in fact, the standard deviation of the proposal distribution, it can be substituted by the production of the CoV and mean of random samples. CoV, which is often indicated as a percentage value and greater than zero, augments generating of random samples by adding an extra level of information to supervise the standard deviation of random field in the proposal distribution. Subsequently, initially, the formulation required to implement the PSAMH method in association with the supported evidence are outlined. Later, a synthetic case delineates the sequence of PSAMH implementation in a comprehensive approach. It should be noted that the presented study is followed by another paper that evaluates the performance of various well-known MCMC samplers in addition to the PSAMH method.

2.3. Methodology

2.3.1. Bayesian and Markov Chain Monte Carlo (MCMC)

Bayes and Price (1763) introduced the conditional probability which later became the backbone of the probabilistic inversion solutions (Vidakovic 1962; Gelman et al. 2013). When the prior and likelihood are assumed to be independent, the posterior distribution is defined as:

$$\pi(\boldsymbol{\theta}|\boldsymbol{d_{obs}}) = \frac{\pi(\boldsymbol{\theta})g(\boldsymbol{d_{obs}}|\boldsymbol{\theta})}{\int \pi(\boldsymbol{\theta})g(\boldsymbol{d_{obs}}|\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{2.1}$$

Where, $\pi(\boldsymbol{\theta})$ and $g(\boldsymbol{d}_{obs}|\boldsymbol{\theta})$ represent the prior and likelihood, respectively. $\boldsymbol{\theta}$ ($\boldsymbol{\theta} = \theta_1, \theta_2, \dots, \theta_d$) denotes the vector of model random variables, and $\boldsymbol{d}_{obs}$ indicates a vector of observed data. $\int \pi(\boldsymbol{\theta})g(\boldsymbol{d}_{obs}|\boldsymbol{\theta})d\boldsymbol{\theta}$ is a normalization constant and marginal distribution of joint distribution of priors and likelihood. Since often in Bayesian analysis, it is required to drive the fraction of two distributions, the constant marginal distribution cancels out from the fraction; and Eq. (2.1) becomes a target distribution as follows:

$$\pi(\boldsymbol{\theta}|\boldsymbol{d}_{obs}) \propto g(\boldsymbol{d}_{obs}|\boldsymbol{\theta}) * \pi(\boldsymbol{\theta}) \tag{2.2}$$

MCMC and MH are well-accepted algorithms when it is either difficult to derive the analytical solution or the posterior is complicated to be sampled. Eq. (2.3) provides the acceptance probability in the MH algorithm.

$$\alpha(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{i-1}) = min\left\{1, \frac{\pi(\boldsymbol{\theta}_i|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_i)}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i-1})}\right\} \tag{2.3}$$

$f(.|.)$ represents the proposal distribution conditioned on the previous step. $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_{i-1}$ denotes the candidate and current samples, respectively. $\alpha(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{i-1})$ provides a probability value indicating how fast the chain moves forward. In addition, MH must satisfy the following condition.

$$ratio = \frac{\alpha(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{i-1})}{U} = \begin{cases} accept & ratio \geq 1 \\ reject & otherwise \end{cases} \tag{2.4}$$

Where, U indicates a randomly drawn number from the Uniform distribution. If the candidate sample is accepted, it stores as the current new sample and the chain moves forward; otherwise, the current sample remains unchanged. Also, the proportion of the accepted samples to the total number of samples determines the acceptance rate. Additionally, random walk metropolis (RWM) defines a special form of MH when the proposal distribution is symmetrical ($f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i-1}) = f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_i)$) so that they cancel out from the MH probability criterion.

$$\alpha'(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{i-1}) = min\left\{1, \frac{\pi(\boldsymbol{\theta}_i|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})}\right\} \tag{2.5}$$

For now, due to the simplification of the computation, it is assumed that the proposal distribution is symmetrical and in particular is the Multivariate Normal distribution (Hoff 2009; Gareth O Roberts et al. 2009). Eq. (2.6) draws a set of random samples at i^{th} iteration.

$$f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i-1}) = \boldsymbol{\theta}_{i-1} + N(\widehat{\mathbf{0}}, \sum) \tag{2.6}$$

Where, $N(.)$ indicates the Multivariate Normal random density function with a zero mean and covariance $\sum$. The function of Multivariate Normal distribution (f_{mvn}) is given as follows:

$$f_{mvn}(\boldsymbol{\theta}_{,j}) = (2\pi)^{-n/2}|\sum|^{-1/2}\exp(-0.5(\boldsymbol{\theta}-\boldsymbol{\mu})^T\sum^{-1}(\boldsymbol{\theta}-\boldsymbol{\mu})) \quad (2.7)$$

Where, $|\sum|$ denotes the determinant of covariance, and μ presents the mean of random parameters. In addition, Eq. (2.8) defines the matrix of covariance.

$$\sum = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_1\sigma_2\rho_{12} \\ \vdots & \ddots & \vdots \\ \sigma_1\sigma_2\rho_{12} & \cdots & \sigma_d^2 \end{bmatrix} \quad (2.8)$$

Where, $\rho_{..}$ indicates the correlation coefficient amidst parameters. In this study, the model random variables are assumed to be independent and identically distributed (i.i.d). Considering the independency of random parameters, $(\rho_{jj'} = 0\ , j \neq j')$ and $(\rho_{jj'} = 1\ , j = j')$, $\sum$ becomes

$$\sum = \begin{bmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_d^2 \end{bmatrix} \quad (2.9)$$

2.3.2. Parallel Scaled Adaptive Metropolis-Hastings (PSAMH)

Authors proposed the Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) algorithm as an augmentation on the adaptive MCMC method that allows automatically optimizing the acceptance rate by constituting several synchronous chains. In order to illustrate how to implement the algorithm, at the current iteration, and using Eq. (2.6), Eq. (2.7) and Eq.(2.9), firstly presume m concurrent chains using the proposal distribution is drawn. Then applying the RWM criterion, Eq.(2.5), the acceptance-rejection state of candidate samples is assessed and accepted samples are stored. Suppose l $(0 \leq l \leq m)$ concurrent chains successfully deliver the RWM criterion. Since it is hypothesized that the model random variables are i.i.d and invoking the definition of RWM, for any two arbitrary chains, k and k', where $1 \leq k\ and\ k' \leq l$, the RWM probability ratios become

$r_{i,k} = \frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})}$ and $r_{i,k'} = \frac{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})}$. Furthermore, the ratio of $r_{i,k}$ and $r_{i,k'}$ approximates by Eq. (2.10).

$$r_{i,k,k'} = \frac{r_{i,k}}{r_{i,k'}} = \frac{\frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})}}{\frac{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})}} = \frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})} \tag{2.10}$$

If $r_{i,k,k'} \geq 1$, then $\boldsymbol{\theta}_{i,k}$ sets as the i^{th} accepted random sample; otherwise, $\boldsymbol{\theta}_{i,k'}$ substitutes by the primary candidate sample. Eq. (2.10) iterates l times to evaluate all accepted samples and eventually identifies the final accepted sample. The above condition implies that the relative probability of occurrence of $\boldsymbol{\theta}_{i,k}$ is higher than $\boldsymbol{\theta}_{i,k'}$ when $r_{i,k,k'} \geq 1$. Therefore, the sampler enables to generate discrete jumps and explore the entire parameter space readily. Note that, the above criterion precisely pursues the RWM acceptance-rejection criterion, hence the detail equilibrium is held; and it is reversible (ergodicity). Moreover, by dropping the term of empirical covariance stemmed in adaptive techniques and accounting on merely the current step sample, the Markovian state of sampling is sustained.

In the general form of MH, when $\hat{r}_{i,k} = \frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i,k}|\boldsymbol{\theta}_{i-1})}$ and $\hat{r}_{i,k'} = \frac{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k'})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i,k'}|\boldsymbol{\theta}_{i-1})}$, Eq. (2.10) is reconstructed as

$$\hat{r}_{i,k,k'} = \frac{\hat{r}_{i,k}}{\hat{r}_{i,k'}} = \frac{\frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i,k}|\boldsymbol{\theta}_{i-1})}}{\frac{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k'})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})f(\boldsymbol{\theta}_{i,k'}|\boldsymbol{\theta}_{i-1})}} = \frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})} \frac{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k})f(\boldsymbol{\theta}_{i,k'}|\boldsymbol{\theta}_{i-1})}{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k'})f(\boldsymbol{\theta}_{i,k}|\boldsymbol{\theta}_{i-1})} \tag{2.11}$$

Eq.(2.11) provides a comprehensive solution when the proposal distribution is asymmetrical. Note that, according to the authors' experience, the rule of thumb for the number of concurrent chains suggested as 5 chains per each random variable; however, it can be different from one case to another.

Additionally, exerting PSAMH method, it is required to approximate the standard deviation of model random variables to derive $\sum$ in Eq.(2.9). That is for, the coefficient of variation (CoV) is used as an operator to supervise the magnitude of the standard deviation. The using of CoV turns to be essential, since most often there is no constructive information about the metric of standard deviations of model random parameters to drive the proposal distribution. Therefore, adopting the idea of tuning the proposal distribution step size by Rosenthal (2011) conveys the problem to the application of coefficient of variation (CoV). By definition, CoV defines the ratio of standard deviation to the mean of samples (Everitt 2006; Forkman and Verrill 2008; Forkman 2009).

$$\mathrm{CoV} = \frac{\sigma}{\mu} \tag{2.12}$$

Where, σ and μ indicate the standard deviation and expected value of generated samples, respectively. Reordering the above equation, the standard deviation computes by Eq.(2.13).

$$\sigma = \mathrm{CoV} * \mu \tag{2.13}$$

By substituting Eq. (2.13) in Eq. (2.9), the general form of the covariance matrix becomes as follows:

$$\sum = \begin{bmatrix} (\mathrm{CoV}_1 * \mu_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathrm{CoV}_d * \mu_d)^2 \end{bmatrix} \tag{2.14}$$

μ updates at each iteration when the chain moves forward. Note that, it is often preferred to set the coefficient of variation equal to a constant number greater than zero ($\mathrm{CoV}_1 = \cdots = \mathrm{CoV}_d = p > 0$). However, when the order of magnitude of random variables is not similar, employing inhomogeneous proposal distribution is recommended (Andrieu and Thoms 2008; Rosenthal 2011). Although, in the case of inhomogeneous proposal distribution, the same principle is applicable regarding the concurrent chains; and despite the analogy of optimum acceptance rate quantities, the algorithm becomes less efficient in comparison to the homogeneous proposal distribution (Gareth O. Roberts and Rosenthal 2002).

It is also necessary to generate random combinations of CoV such that they fulfill the optimum acceptance rate criterion. Forkman and Verrill (2008) and Forkman (2009) suggested the Chi square distribution as an approximation of CoV distribution for the Normal proposal distribution when the ratio of σ / μ becomes less than 1/3 ($CoV = chi2(1.5)$). Fig. 2.1 demonstrates an example of the combinations of CoV when the random field size is three.

Additionally, take into consideration that, often one requires applying constraints on the generated random samples through the Normal proposal distribution. The truncated Normal distribution (Robert 1995; Burkardt 2014) is an alternative which enables the practitioner to take advantage of general configuration of Normal distribution and supervise generated samples. Eq.(2.15) presents the two-sided truncated Normal distribution. The significance of mentioned equation is in its application in Eq. (2.11) due to asymmetrical shape of the proposal distribution.

$$f\left(\boldsymbol{\theta}\middle|\boldsymbol{\mu},\boldsymbol{\theta}_{lower},\boldsymbol{\theta}_{upper},\sigma\right)=\frac{\exp(-(\boldsymbol{\theta}-\boldsymbol{\mu})^2/2\sigma^2)}{\sqrt{2\pi}\sigma(\Phi\left(\frac{\boldsymbol{\theta}_{upper}-\boldsymbol{\mu}}{\sigma}\right)-\Phi\left(\frac{\boldsymbol{\theta}_{lower}-\boldsymbol{\mu}}{\sigma}\right))} \tag{2.15}$$

Where, Φ denotes the standard cumulative Normal distribution. $\boldsymbol{\theta}_{upper}$ and $\boldsymbol{\theta}_{lower}$ represent the constraint of Normal proposal distribution, respectively.

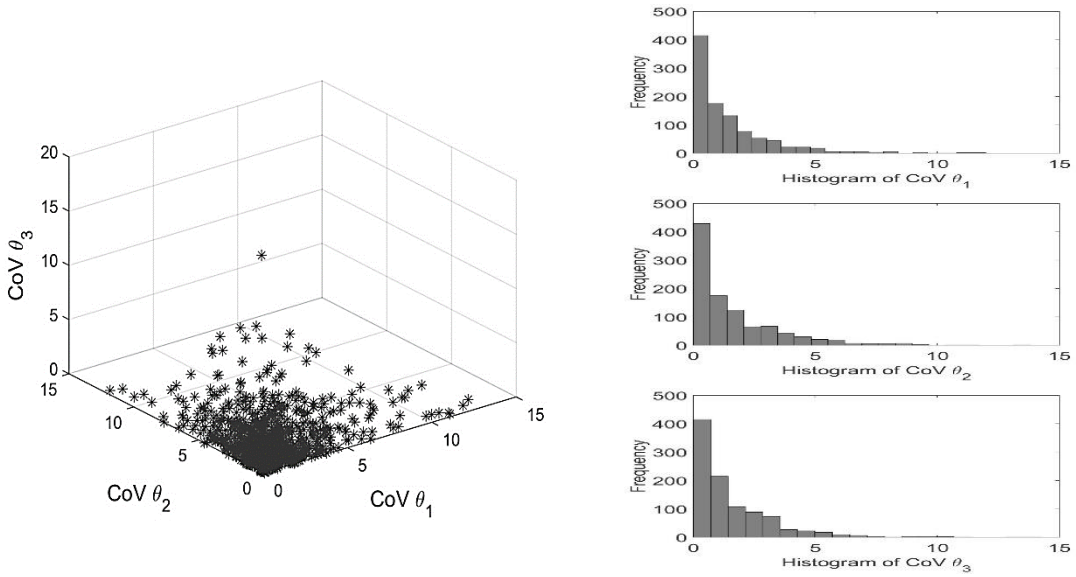

Fig. 2.1. Left, typical configuration of 1000 combinations of CoV, and right, histogram of CoV corresponding with each random variable

Exerting Eq. (2.15) in the second term of the general form of MH criterion in Eq.(2.3), the subsequent formula is obtained.

$$\frac{f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i-1})}{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_i)}=\frac{\sqrt{2\pi}\sigma\,\exp(-(\boldsymbol{\theta}_i-\boldsymbol{\theta}_{i-1})^2/2\sigma^2)\,(\Phi\left(\frac{\boldsymbol{\theta}_{upper}-\boldsymbol{\theta}_i}{\sigma}\right)-\Phi\left(\frac{\boldsymbol{\theta}_{lower}-\boldsymbol{\theta}_i}{\sigma}\right))}{\sqrt{2\pi}\sigma\,\exp(-(\boldsymbol{\theta}_{i-1}-\boldsymbol{\theta}_i)^2/2\sigma^2)\,(\Phi\left(\frac{\boldsymbol{\theta}_{upper}-\boldsymbol{\theta}_{i-1}}{\sigma}\right)-\Phi\left(\frac{\boldsymbol{\theta}_{lower}-\boldsymbol{\theta}_{i-1}}{\sigma}\right))} \tag{2.16}$$

By equality of $\exp(-(\boldsymbol{\theta}_i-\boldsymbol{\theta}_{i-1})^2/2\sigma^2)=\exp(-(\boldsymbol{\theta}_{i-1}-\boldsymbol{\theta}_i)^2/2\sigma^2)$, Eq. (2.16) becomes:

$$\frac{f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i-1})}{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_i)} = \frac{\Phi\left(\frac{\boldsymbol{\theta}_{upper}-\boldsymbol{\theta}_i}{\sigma}\right) - \Phi\left(\frac{\boldsymbol{\theta}_{lower}-\boldsymbol{\theta}_i}{\sigma}\right)}{\Phi\left(\frac{\boldsymbol{\theta}_{upper}-\boldsymbol{\theta}_{i-1}}{\sigma}\right) - \Phi\left(\frac{\boldsymbol{\theta}_{lower}-\boldsymbol{\theta}_{i-1}}{\sigma}\right)} \tag{2.17}$$

Henceforth, when it becomes necessary, the second term of MH criterion substitutes by Eq. (2.17).

2.3.3. Acceptance rate

It is observed that when the step size of proposal distribution gets a noticeably small value, chain subsequently takes a considerable long time to move forward. In contrast, a large step size induces more sample rejection, and hence, the chain will not progress. Gelman et al. (1996), Roberts et al. (1997), and Bédard (2007) separately investigated the optimized acceptance rate for RWM and recommended that the optimal number when the parameter space size (d) tends to infinity is 23.4% and when d is unit, acceptance rate becomes about 44%. Gelman et al. (1996) specifically tabulated the correlation amid the random field size, d, and the acceptance rate up to d=10. Fig. 2.2 demonstrates the same data and a proposed hyperbolic fit curvature applying the non-linear least square method (The parameter space size is set up to d=100) to be able to approximate the optimum acceptance rate when the space dimension varies.

Eq. (2.18), which is retrieved from the optimization analysis, develops an analytical solution to compute the associated acceptance rate where the dimension size is other than already computed.

$$\xi_{target}(d) = 0.234 + \frac{0.654}{(1+1.4307d)^{1/0.775}} \tag{2.18}$$

Moreover, in the absence of any preferred sampling method, it is generally recommended that the acceptance rate should not be less than 15% or larger than 50% (Gareth O. Roberts and Rosenthal 2002; Rosenthal 2011).

2.4. PSAMH augmentation features

In this section, several possible features are introduced that allow to expedite the MCMC convergence and optimize the running time.

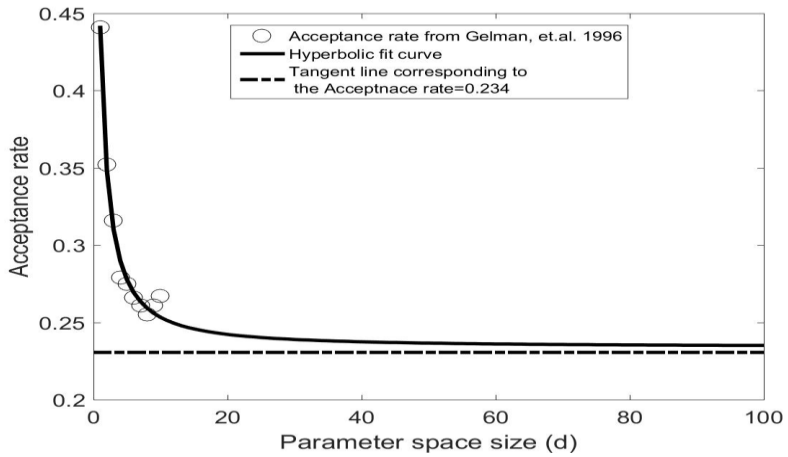

Fig. 2.2. Hyperbolic curvature fits to the acceptance rate corresponding to the parameter space size

2.4.1. Scaling factor (γ)

The scaling factor defines a multiplication term to optimize the acceptance rate by fluctuating the proposal step size. Although, the scaling factor induces to the proposal distribution after a fixed number of iterations (batches), assigning the batch size appears as an essential step that should be dealt with wisely. In order to reduce the unnecessary alteration impact on the generated samples' process, it is advisable to assign a fairly small value to the batch size on the course of early stage of sampling (such as 50 or 100

iterations); and then after a specific iteration (for instance, the Burn-In point) it is increased to a larger quantity (such as 500 or 1000 iterations). Haario et al. (2001) and Roberts & Rosenthal (2009) introduced the scaling term and how to apply it over the computation procedure. The scale factor, γ, invariably begins with the unit value. Then, recurrently, after a certain number of iterations (batch size) the state of the current acceptance rate, ξ_i, compares with the optimum acceptance rate, ξ_{target} via Eq. (2.19).

$$\gamma_i = \begin{cases} \gamma_{i-1} + \varepsilon & \xi_i > \xi_{target} \\ \gamma_{i-1} - \varepsilon & \xi_i < \xi_{target} \end{cases} \tag{2.19}$$

Where, γ_i denotes the scaling factor at the current iteration, i. ξ_i and ξ_{target} indicate the current and optimum acceptance rate. The quantity of ε should be assigned deterministically; and it can be obtained by trials and errors aimed to make ξ_i becomes as close as possible to the optimum acceptance rate. Nevertheless, in this study, Eq. (2.20) is suggested as another alternative which admits the same result but removes the necessity of manually tuning the value of ε.

$$\gamma_i = \xi_i / \xi_{target} \tag{2.20}$$

If the current acceptance rate is larger than the target, γ_i becomes larger than one, hence, enlarges the proposal step size. When the current acceptance rate, in contrary, turns to be smaller than the target, the step size consequently becomes smaller. Moreover, Eq. (2.20) is implicitly analogous to the gradient of the current and target acceptance rate. It also connotes that the order of magnitude of scaling factor automatically adjusts by taking apart or approaching to the true mean of the process. Note that, in the PSAMH framework

the scale factor applies to the entire matrix of step size to shift them to the optimum acceptance rate through Eq. (2.21).

$$\sum = {\gamma_i}^2 * \begin{bmatrix} (\mathrm{CoV}_1 * \mu_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathrm{CoV}_d * \mu_d)^2 \end{bmatrix} \tag{2.21}$$

2.4.2. Reduction of redundancy in concurrent chains

Earlier, it is mentioned that, m concurrent chains with distinct combination of CoV should be drawn to construct MCMC. Considering the histogram plot of CoV in Fig. 2.3, left, it is observed that some combinations of CoV are invoked more frequently in compare to others. Therefore, using the histogram function allows to keep more frequent applied combinations and remove the redundant CoV after a fix number of iteration (Fig. 2.3, right).

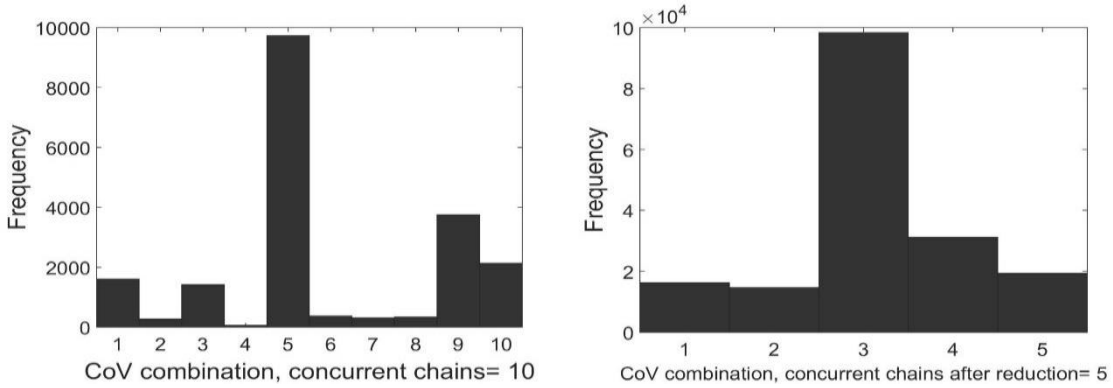

Fig. 2.3. Frequency of selection of CoV, left, 10 concurrent chains and right, after reduction to 5 chains

In order to identify the number of iteration which after that the reduction mechanism is applicable, it becomes, hence, necessary to diagnose the state of stationary. Monitoring the online mean and standard deviation (also known as the cumulative mean and standard deviation) of random variables develops the notion of recognition the MCMC stationary

condition. Acquiring the stationary condition allows to set the burn-in point, which indicates the number of iteration denoting the stationary condition of experiment. When corresponding plots of the online mean and standard deviation turn to straight lines, it can be interpreted as achieving the stationary condition (see Fig. 2.6 for examples). Note that, the reduction of redundant CoV is mainly important when the number of synchronous chains is initially large, and reducing the number of chains plays a key role to optimize the experiment running time.

2.4.3. Adaption mechanism vanishing after stationary condition

Intuitively, it is preferred to remove any adaption mechanism and continue the chain merely relying on the regular MH criterion. Andrieu & Thoms (2008) described a formulation to vanish the adaption. However, it can also be implemented by simply removing any adaption mechanism after the stationary condition.

2.4.4. Employing empirical covariance

In analogy to AMCMC methods, it is plausible to induce the correlation structure between random parameters through the empirical covariance. However, applying the empirical covariance requires investigating the state of ergodicity in MCMC for each proposal distribution and yet would not guaranty the convergence to the true posterior and also removes the Markovian state of the experiment.

2.4.5. Using parallel computing toolbox

Considering a circumstance that the forward model is substantially sophisticated or robust, it is useful to split the concurrent chains into several single tasks to be able to

assign them to separate physical computer cores (workers). Applying the parallel toolbox of modeling software, the system technically breaks the chain and sends the information to the workers. Workers, accomplishing the assigned tasks, return the results to the main system and the procedure continues to the last chain. The main system, after assembling all information from workers, performs the final computation, reserves the result and proceeds to the next step. Note that the procedure of sending and receiving information is relatively time consuming; hence, this feature is not recommended for simple forward models.

2.5. PSAMH implementation pseudo procedure

The concise implementation sequence of PSAMH framework is recapitulated subsequently.

- Set the initial values for random variables, such that they provide a non-zero target distribution function results. A recommended mechanism to obtain appropriate initial values is discussed later in the case study one.
- Set the number of concurrent chains, m, and the batch size.
- Approximate the optimum acceptance rate using Eq. (2.18)
- Generate once $CoV_j = [1, d], (j \in [1, m])$ from Chi square distribution ($Chi2(1.5)$) and construct $\sum_{i,j}$ applying Eq. (2.14)
- Draw m concurrent candidate samples from the proposal distribution using Eq. (2.6) and Eq.(2.7). Also, supply the truncated normal distribution from Eq.(2.17) in the case of constraint proposal distribution

- Accept candidate samples by probability of α and reject them by $1-\alpha$ (Eq. (2.3) or Eq.(2.5) and Eq.(2.4)) and store l accepted samples
- Identify the final accepted candidate by inserting l accepted samples into Eq. (2.10) or Eq.(2.11)
- Compare the current acceptance rate, ξ_i, with the optimum acceptance, ξ_{target}, for the iteration equals the batch size and update the scale factor applying Eq. (2.20) and Eq. (2.21)
- Remove the redundant CoV combinations after the stationary condition achieved
- Store the final accepted sample and proceed to the next iteration

Fig. 2.4 displays the diagram of PSAMH framework in order to provide a visual perception of implementation sequences.

2.6. Experimental design

A numerical synthetic case elaborates the implementation sequence of a stochastic process in MCMC using the PSAMH framework. It is assumed that an experiment with 10 fix locations which indicates the time or displacement of the process constitutes a set of observation.

Table 2.1 exhibits three case studies that is constructed by varying the repetition of observations at fix locations. The use of repetition aims to provide a tool to assess the influence of increasing the number of observations applying PSAMH method.

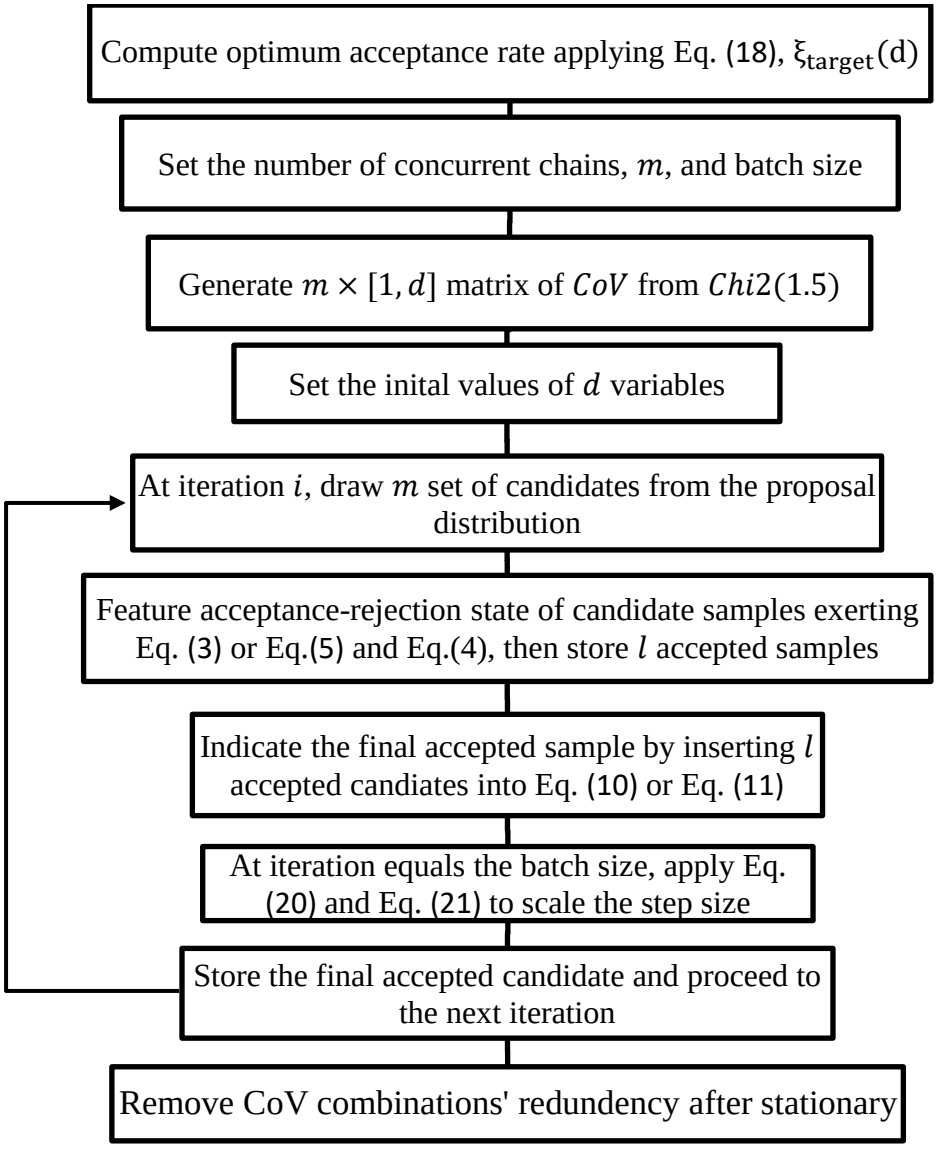

Fig. 2.4. PSAMH implementation diagram

Table 2.1 Experimental design for the stochastic process

Case	Locations	Repeats	Total number of observation
P10R1	10	1	10
P10R5	10	5	50
P10R10	10	10	100

2.7. Stochastic process

Eq. (2.22) serves as the forward model function to represent the negative exponential behavior of the observation.

$$f(x) = Ie^{-Cx} \tag{2.22}$$

Where, I and C are the intercept and curvature coefficients of the exponential function and they are respectively set to 1 and 0.4. The synthetic observations, which are randomly generated applying the Normal distribution with the expected values retrieved from the forward model, is obtained from Eq. (2.23) and demonstrated in Fig. 2.5.

$$d_{observe}(x) \sim N(f(x), \sigma_{likelihood} = 0.1) \tag{2.23}$$

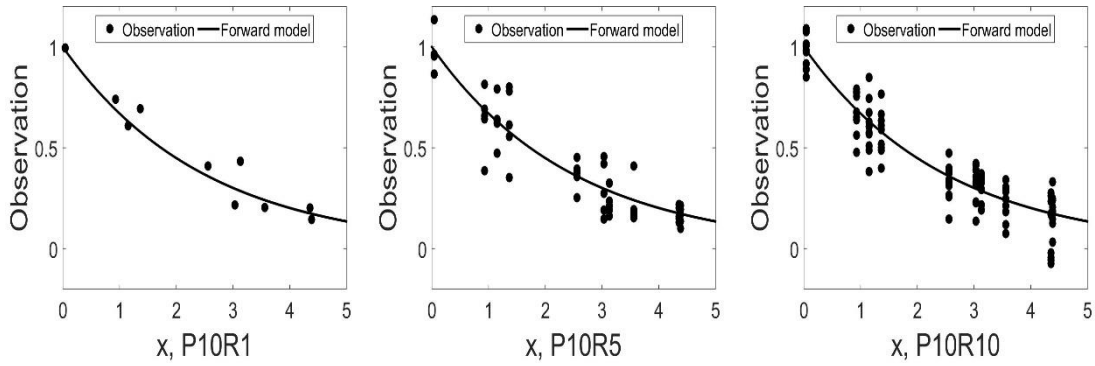

Fig. 2.5. Observation and exponential curve corresponding to Table 2.1, for cases from left to right P10R1, P10R5 and P10R10

In this experiment, I, C and the standard deviation of likelihood ($\sigma_{likelihood}$) indicate the random variables in the MCMC stochastic analysis. The priors are assumed to be i.i.d. and non-informative. Following, procedures required to implement the Bayesian process applying PSAMH framework are elaborated.

2.7.1. Optimization

Often there is no initial information about the statistics of likelihood or initial values of random variables. The least square optimization method initiates an alternative to provide the preliminary information for MCMC. The residual (error) between the least square fit and observation data develops an insight about the distribution of likelihood whilst the optimized values serve as the starting points in MCMC. In this experiment, it is hypothesized that the distribution of error projects the Normal distribution.

2.7.2. Sampling using PSAMH algorithm

Approximating the distribution of likelihood and initial values, and by applying PSAMH method, allow to draw m candidate samples from the proposal distribution. Additionally, two features of scaling factor and reduction of chains redundancy are employed to augment the sampling mechanism. The initial number of concurrent chains is set to 10 ($m = 10$); however, after 40,000 iterations (stationary condition) it is reduced to 5 chains. The MCMC lasted for 200,000 iterations to provide enough samples after stationary. The burn-in point gets the value of 40,000, which is concluded from the online mean and standard deviation plots. MCMC experiments of three experimental designs exerting the PSAMH algorithm and online mean and standard deviation associated with random variables are demonstrated in Fig. 2.6. Furthermore, the fixed scale of plots aimed to readily compare experiments in one plot.

Considering Fig. 2.6, it is realized that by increasing the number of observations the MCMC experiment presents less dissemination. The same behavior can be concluded from the plots of online mean and standard deviation. Comparing P10R1, P10R5, and

P10R10, it is implied that while the online mean almost constantly approximates the same value, standard deviation depicts smaller values by adding more data progressively.

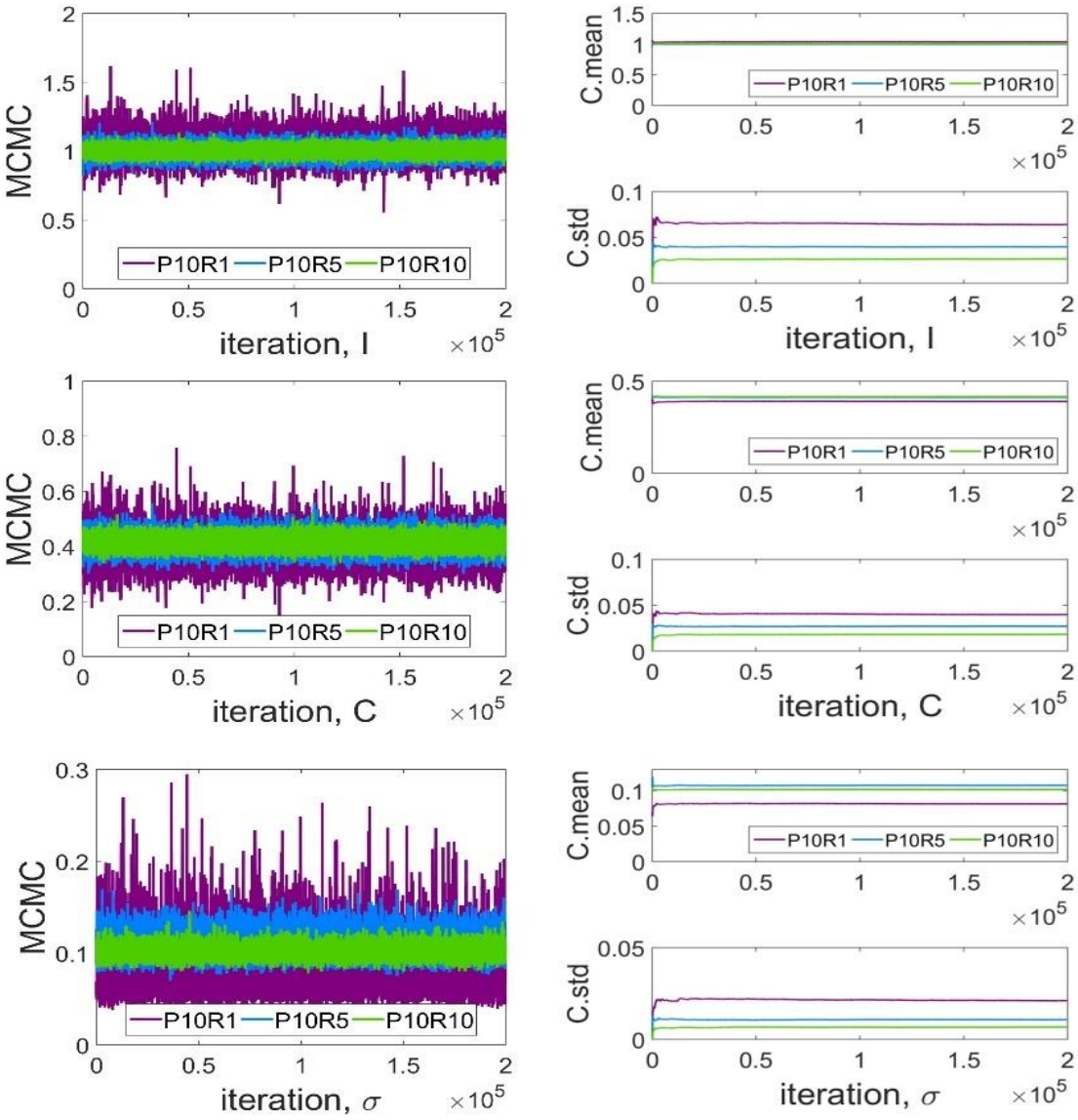

Fig. 2.6. Left, MCMC experiments, and right, online mean and standard deviation of random variables

Table 2.2 illustrates the associated input data, constraints of random variables, and Bayesian inferences after computation. The considerable wide ranges of priors' constraints using the Uniform distribution allow to avoid the influence of limitations on

sampling. Also, the likelihood of target distribution sets as the Normal distribution with unknown standard deviation ($\sigma_{likelihood}$) alluded from the Optimization section.

Table 2.2 Input values and statistics of experiments

Variable	Parameter Range [min, max]	Case	True Value	Mean	Mode	Standard Deviation	5% Percentile	95% Percentile
I	[1e-5,9.99]	P10R1	1	1.036	1.055	0.064	0.94	1.140
		P10R5		0.997	0.998	0.039	0.932	1.063
		P10R10		1.002	1.021	0.027	0.959	1.046
C	[1e-5,9.99]	P10R1	0.4	0.389	0.397	0.039	0.328	0.456
		P10R5		0.412	0.418	0.027	0.369	0.458
		P10R10		0.416	0.417	0.018	0.386	0.446
$\sigma_{likelihood}$	[1e-5,1e2]	P10R1	0.1	0.081	0.065	0.021	0.055	0.121
		P10R5		0.108	0.109	0.011	0.092	0.127
		P10R10		0.102	0.092	0.007	0.091	0.113

Retrieved means and modes of random variables connote the fact that by increasing the number of observations at each location the precision of computed results improves progressively. In accordance, by souring the number of observations the standard deviation of each variable admits a similar conclusion by producing smaller values. The same state holds for 5% and 95% percentiles, as they provide smaller ranges by increasing the number of observations.

2.7.3. Statistics

In addition to the first and second order of statistics presented in Table 2.2, the relative frequency histogram, joint density and cumulative distribution function also deliver the statistics of random variables.

The influence of updating observations is better depicted in the relative frequency histogram and cumulative density function plots (Fig. 2.7). The tip of the histogram plot

gradually elevates from P10R1 to P10R10; whereas, the dispersion in data abates. Moreover, $\sigma_{likelihood}$ of P10R1 presents some degree of right skewness as it is expected for the standard deviation.

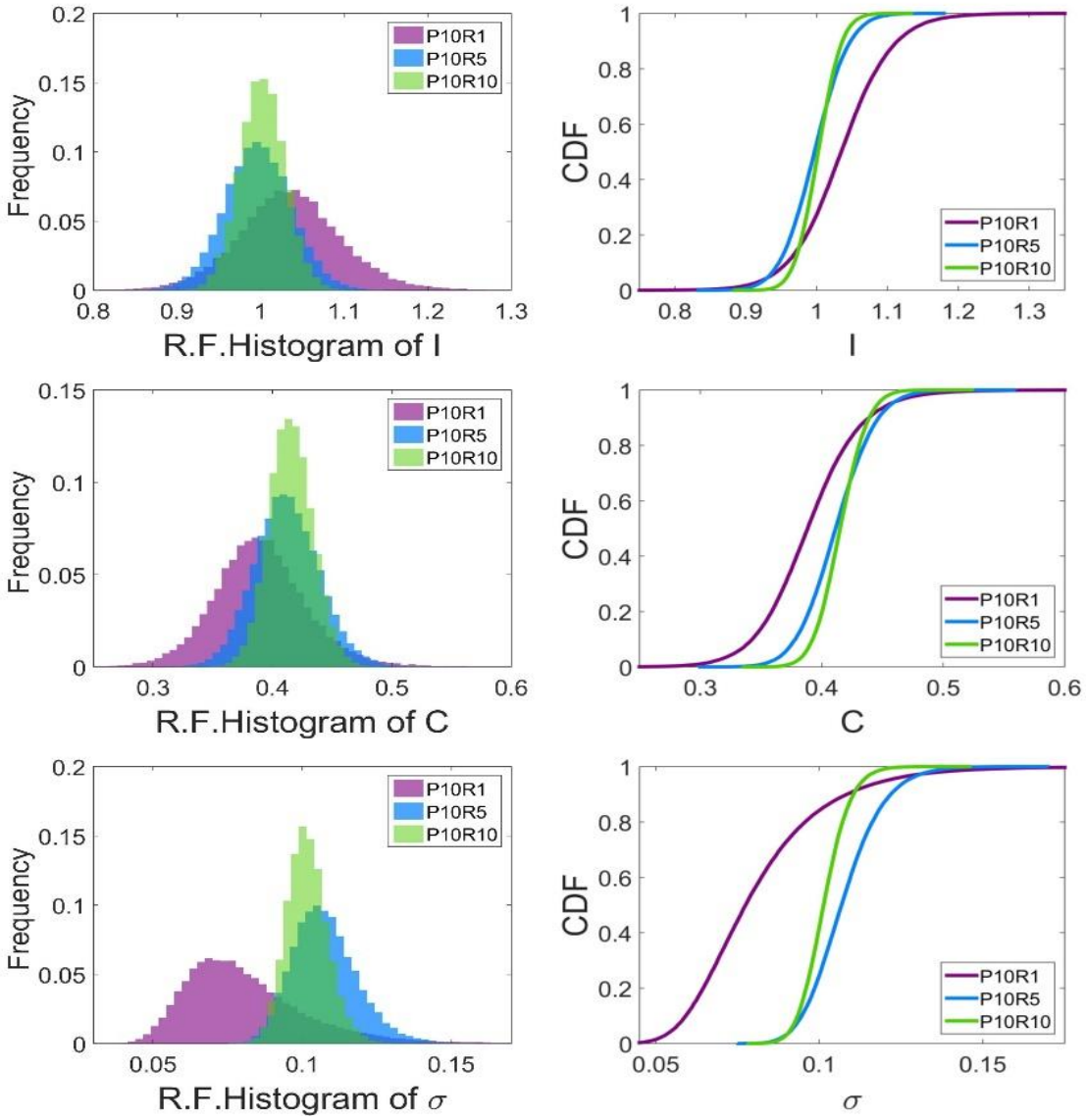

Fig. 2.7. Left, relative frequency histogram, and right, cumulative density function of experiments

It is well evident that the standard deviation of random samples approximates Chi square distribution when samples express the Normal behavior. However, when the

number of observation increases, the configuration of histogram and CDF plots asymptotically approximate the central limit theorem, hence, presenting some degrees of Normality.

Fig. 2.8 demonstrates the joint frequency histogram of variables I and C regarding all numerical experiments. Histograms are plotted in the same scale to be able to compare them. Plots of joint density histogram clearly exhibit the influence of updating observations on increasing the confidence over the analysis. Fig. 2.8 also illustrates the positive correlation structure among the random variables I and C. Inferences retrieved from the vector of posterior have become a common practice to predict the characteristics of random field, especially when the information regarding one random variable is unreliable or unachievable, whereas the correlation coefficient and statistics with respect to the other conjugated random variable are known.

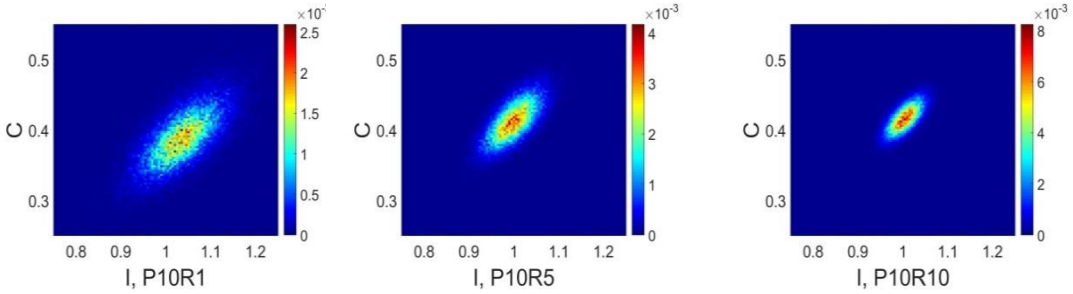

Fig. 2.8. Comparison between the joint frequency histogram of case studies

2.7.4. Model realizations

Bayesian paradigm constructs a full access to not only the posterior space, but also the event space through model realizations. Model realization features an applicable tool to pass thousands of realizations through the observation by randomly drawing thousands of pair of samples from the posterior vector and inserting them into the forward model.

Thus, it becomes possible to retrieve the inferences such as mean and standard deviation of model realizations. Note that, the standard deviation of the model realizations often serves as a reliable gauge to assess the certainty of analysis.

Fig. 2.9 and Fig. 2.10 illustrate 10,000 realizations together with the mean and standard deviation of model realizations.

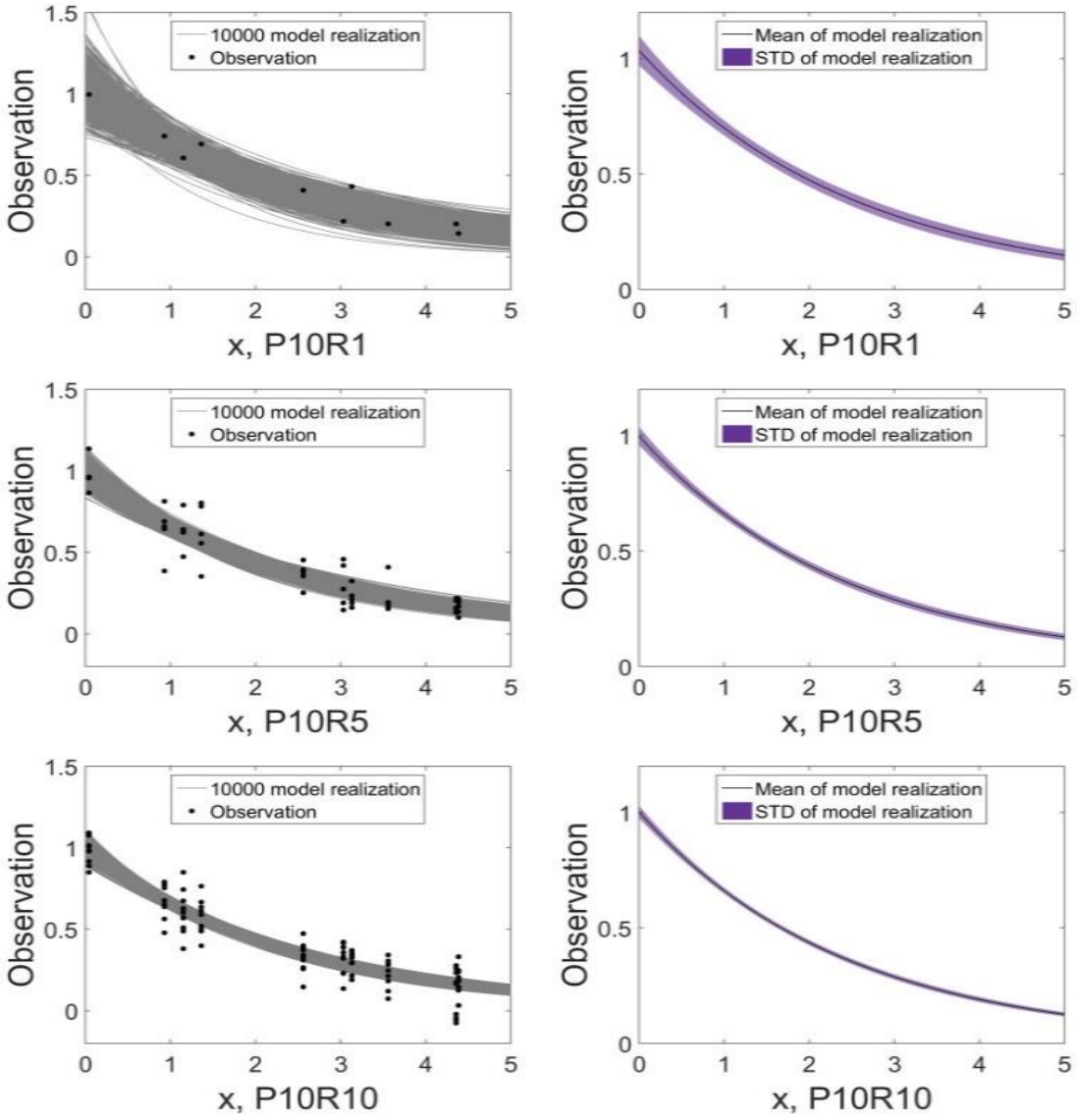

Fig. 2.9. Top, 10,000 realizations and observation, bottom, mean and standard deviation of realization

Also, the disseminations of 10,000 realizations coupled with the observation are shown in Fig. 2.9 left. Fig. 2.9 right exhibits the diffusion of realizations around the mean. By moving from up to down, it is observed that the level of confidence increases by adding more data. In all cases, the uncertainty from the analysis appears higher for the smaller x and decreases by progressing on the course of increasing x.

Fig. 2.10, additionally, depicts a comparison between mean and standard deviation of 10,000 realizations for experiments together.

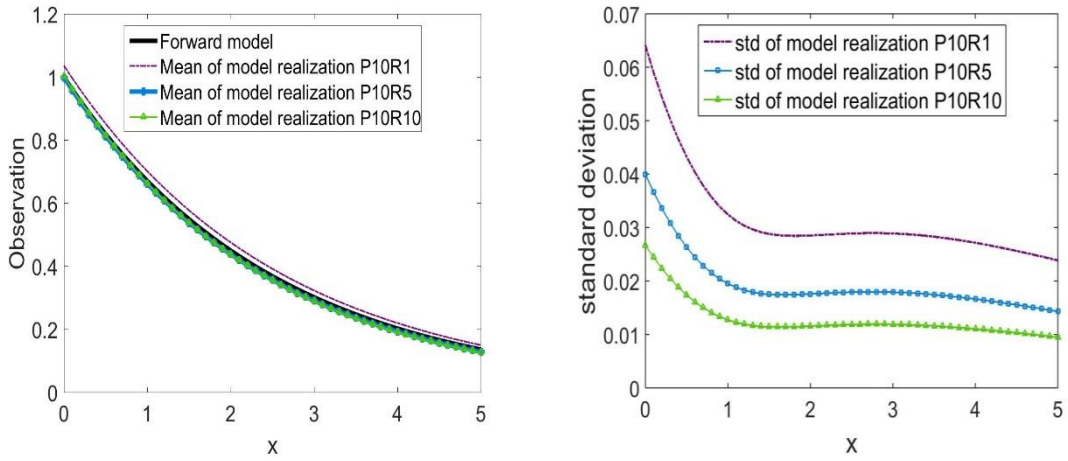

Fig. 2.10. Comparison between left, means and right, standard deviations of model realization of experiments and forward model

Mean of realizations, left, displays the significance of obtaining more data when the mean of model realization plots of cases P10R5 and P10R10 are coincided and close to the forward model curvatures. The same trend can be better perceived from the standard deviation comparison plots where the plot of P10R10 lies beneath all cases.

2.7.5. Acceptance rate and scale factor

Fig. 2.11 depicts a typical trend of acceptance rate and scale factor for 2e5 iterations after the convergence is achieved. It is earlier mentioned that one option to determine the burn-in point (stationary condition) is examining plots of online mean and standard deviation, and in the current case study by assessing the associated plots the burn-in point is set to 40,000.

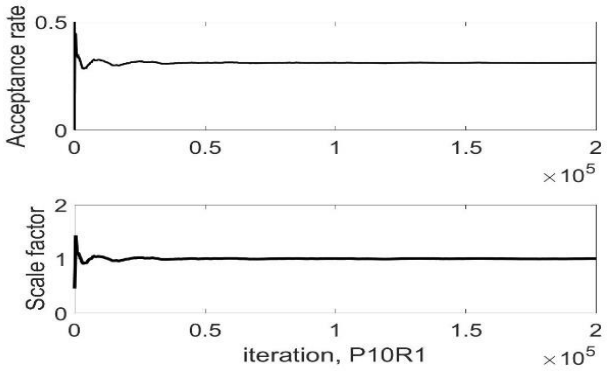

Fig. 2.11. Acceptance rate and scaling factor for P10R1

Plots of acceptance rate and scale factor as another alternative are proposed to evaluate the stationary condition of experiment. Both plots demonstrate a constant trend after almost 40,000 iterations which are in compliance to the chosen burn-in point. Some concise remarks of conclusions are provided in section 8.1.1.

3. PRACTICAL BAYESIAN FRAMEWORK TO PARALLELIZE ADAPTIVE MARKOV CHAIN MONTE CARLO: PERFORMANCE ASSESSMENT

3.1. Overview

Several Markov Chain Monte Carlo (MCMC) techniques have been recently developed which couple the acceptance-rejection criterion of Metropolis-Hastings (MH) algorithm with either the parallel mechanism of synchronous independent chains or tuning the proposal distribution of MH to fully integrate the random space. This study, by assessing the performance of several robust MCMC methods, recognizes the implementation sequence of a novel approach that is called the Parallel Scalded Adaptive Metropolis-Hastings (PSAMH), which is introduced in the part one of this study. The use of a synthetic experiment, additionally, allows the practitioner to comprehensively percept the discrepancy amidst driven MCMC methods in order to eventually appreciate the more applicable sampler. Random Walk Metropolis (RWM), Parallel Tempering (PT), Multiple-Tries Metropolis (MTM), Adaptive Metropolis (AM) and so called Scaled Adaptive Metropolis (SAM) are other chosen frameworks to constitute the performance evaluation. Also, altering the standard deviation and correlation coefficient of the proposed target distribution is the key technique to incorporate the synthetic experiment. The results, by advocating the functionality of all samplers, finally imply that methods of PSAMH and then PT perform better where the problem becomes substantially sophisticated.

3.2. Introduction

This work develops a complement study of a recently introduced MCMC sampling method which is called the Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) in section 2 and aims to provide an insight regarding the performance of this method when compared with several well-known samplers.

Recent significant developments in the computational technology have made scholars able to constitute more robust techniques dealing with Bayesian inferences. Despite the irrefutable privileges have been provided by constructed frameworks, practitioners inevitably committed to a higher level of perplexity in selecting the appropriate probabilistic method aimed to perform better concerning their specific problems. That is for, in this study, the implementation mechanism of several Markov Chain Monte Carlo samplers (MCMC) (Gilks and Richardson 1996; Geyer 2002; Faming, Chuanhai, and Raymond 2010; Berg and Billoire 2008) in addition to PSAMH are elaborated to delineate the performance of each method when some contributing factors are varied.

Bayesian theorem applying MCMC together with the Metropolis-Hastings (MH) (Metropolis et al. 1953; Hastings 1970) algorithm permits integrating random samples using the acceptance-rejection criterion of MH to eventually converge to the true mean of the posterior space. Besides, the proportion of accepted samples to the total number of generated samples constitutes the acceptance rate of sampling. Typically, MCMC samplers applying the prior evidence associated with the random field and accurate physical models (also known as forward model) are formulated to either tune the proposal distribution of MH algorithm or alter the target distribution (Kendall, Liang, and Wang

2005a). Subsequently, several MCMC samplers are introduced that share the application of MH algorithm in their structures.

Random Walk Metropolis (RWM) (Hoff 2009; Gelman et al. 2013) denotes a common MCMC sampler when the proposal distribution of MH algorithm becomes symmetrical. RWM as a basic algorithm defines an essential baseline to be able to compare the performance of other samplers.

In addition, Adaptive Markov Chain Monte Carlo (AMCMC) (Sims 1998; Andrieu and Thoms 2008; Atchadé et al. 2009; Rosenthal 2011; Graves 2011) and Adaptive Metropolis (AM) (Haario, Saksman, and Tamminen 1999; Haario, Saksman, and Tamminen 2001) induces the correlation structure between random variables applying the empirical covariance. The step size of proposal distribution in AM algorithm requires to be tuned deterministically to optimize the acceptance rate. It is evident that the search becomes optimized when the space size varies between one to infinity, the acceptance rate fluctuates from 0.44 to 0.234, respectively (Gelman, Roberts, and Gilks 1996; G. O. Roberts, Gelman, and Gilks 1997). Furthermore, Haario et al. (2001),; Andrieu and Thoms (2008); and Roberts and Rosenthal (2009), defined a scaling factor that by multiplying to the AM main formula enables the sampler to achieve the optimum acceptance rate, that in this study is so called Scaled Adaptive Metropolis (SAM). Craiu et al. (2009), besides, aggregated the method of AMCMC with parallelizing chains and invoking the concurrent chains after a fix number of iterations (batch size).

Author, as it mentioned earlier, introduced PSAMH as a robust framework that constitutes synchronous MCMC chains merged with the Adaptive MCMC methodology.

PSAMH, hence, automatically tunes the step size of MH proposal distribution aims to optimize the acceptance rate of posterior space. The formulation and implementation sequence of PSAMH algorithm is comprehensively elaborated in the first part of the presented series works.

Another method which constitutively relies on the MH algorithm is the Parallel Tempering framework (PT) (Hukushima and Nemoto 1996; Wang and Swendsen 2004; Earl and Deem 2005). PT generates several parallel chains while tempering the target distribution. This method is constructed applying the concept of traveling from a chain to another in order to reduce the level of energy among the parallel chains. It is initially hypothesized that chains tend to be stabilized by moving from the higher to lower energy state. Miasojedow et al. (2013), Calderhead (2014) and Cotter et al. (2015) assembled the PT and AMCMC and proposed robust algorithms to contribute the influence of both methods. In addition, Calderhead (2014) described a general approach toward parallelizing Adaptive Metropolis framework (AM).

Multiple-Tries Metropolis (MTM) (Liu, Liang, and Wong 2000; Craiu and Lemieux 2007) is another alternative technique to explore the posterior space by also applying the MH algorithm. Multiple chains that run synchronously construct the elements of weighted acceptance probability criterion. Yang et al. (2016) recently assessed the effect of merging the adaptive sampling into MTM and exhibited some degrees of improvements in the efficiency of MTM. The privilege of applying MTM or PT is embedded in their independency of selecting appropriate MCMC starting values and generating less dependent random samples. However, implementation of these methods requires a higher

level of probabilistic and statistic expertise when several strata of acceptance-rejection criteria should be constructed. Subsequently, the implementation sequences required to drive RWM, PSAMH, AM, SAM, PT and MTM sampling techniques in association with a synthetic case study are outlined.

3.3. Methodology

Subsequently, in addition to PSAMH, several MCMC samplers' techniques that are employed to evaluate their performance are elaborated.

3.3.1. Random Walk Metropolis (RWM)

RWM sampler signifies a special form of MH when the proposal distribution is symmetrical, hence ($f(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{i-1}) = f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_i)$), therefore, they cancel out from the MH criterion.

$$\alpha(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{i-1}) = min\left\{1, \frac{\pi(\boldsymbol{\theta}_i|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i-1}|\boldsymbol{d}_{obs})}\right\} \tag{3.1}$$

RWM exhibits the core technique of MH sampling also exerted in other samplers and is delineated by the following steps.

- Set the initiate quantities of random variables
- Draw candidate sample applying Eq.(2.6)
- Tune the step size (with trial and error) such that the acceptance rate approaches to the optimum acceptance rate (Eq.(2.18))
- Investigate the acceptance state of candidate samples with the RWM criterion using Eq.(3.1)
- Substitute the current with accepted samples

- Proceed to the next iteration

3.3.2. Parallel Scaled Adaptive Metropolis-Hastings (PSAMH)

PSAMH, technically, draws several synchronous chains from a proposal distribution to automatically optimize the acceptance rate of MH criterion. The use of a scaling mechanism, in addition, allows to fluctuate the step size of proposal distribution. The implementation sequence of PSAMH framework is concisely elaborated subsequently.

- Set the initial values for random variables, such that they provide a non-zero target distribution function results
- Set the number of concurrent chains, m, and the batch size
- Generate once $CoV_j = [1, d], (j \in [1, m])$ from Chi square distribution ($Chi2(1.5)$) and construct $\sum_{i,j}$ applying Eq.(3.2)

$$\sum_{i,j} = \begin{bmatrix} (\mathrm{CoV}_1 * \theta_{i-1,1})^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathrm{CoV}_d * \theta_{i-1,d})^2 \end{bmatrix} \tag{3.2}$$

- Draw m concurrent candidate samples from the proposal distribution using Eq. (2.6)
- Accept candidate samples by probability of α and reject them by $1 - \alpha$ (Eq. (2.3)) and store l accepted samples
- For any two arbitrary chains, k and k', where $1 \leq k \; and \; k' \leq l$, identify the final accepted candidate by inserting l accepted samples into Eq.(3.3)

$$\hat{r}_{i,k,k'} = \frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})} \frac{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k}) f(\boldsymbol{\theta}_{i,k'}|\boldsymbol{\theta}_{i-1})}{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k'}) f(\boldsymbol{\theta}_{i,k}|\boldsymbol{\theta}_{i-1})} \tag{3.3}$$

Where, $r_{i,k,k'}$, specifies the ratio of two accepted samples with respect to the current sample. Furthermore, Eq.(2.4) permits evaluating the state of final accepted sample retrieved ratio from Eq.(3.3)

$$accepted\ sample = \boldsymbol{\theta}_i = \begin{cases} \boldsymbol{\theta}_{i,k} & \hat{r}_{i,k,k'} \geq 1 \\ \boldsymbol{\theta}_{i,k'} & \hat{r}_{i,k,k'} < 1 \end{cases} \tag{3.4}$$

Eq. (3.3) and Eq. (2.4), together, iterate l times to assess all accepted samples and eventually identify the final accepted sample.

- At iteration equals to the batch size, recurrently, compare the current acceptance rate, ξ_i, with the optimum acceptance, ξ_{target}, and update the scale factor applying Eq.(3.5) and Eq. (3.6)

$$\gamma_i = \xi_i / \xi_{target} \tag{3.5}$$

$$\sum = {\gamma_i}^2 * \begin{bmatrix} (\mathrm{CoV}_1 * \theta_{i-1,1})^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathrm{CoV}_d * \theta_{i-1,d})^2 \end{bmatrix} \tag{3.6}$$

Take into consideration that, the scale factor, originally is a vector contains value one and merely updates at the batch size iteration.

- Store the final accepted sample and proceed to the next iteration

3.3.3. Adaptive Metropolis (AM)

Gareth O. Roberts and Rosenthal (2009) denoted the proposal distribution of AM as follows.

$$f(\theta_i|\beta, d, \sum_i) = \begin{cases} N(\theta_{i-1}, (0.1)^2 I_d/d) & i \leq 2d \\ (1-\beta)N\left(\theta_{i-1}, \frac{(2.38)^2 \sum_i}{d}\right) + \beta N\left(\theta_{i-1}, \frac{(0.1)^2 I_d}{d}\right) & i < 2d \end{cases} \quad (3.7)$$

Where, d indicates the parameter space size, $\sum_i$ presents the empirical covariance matrix for $i-1$ iterations and β identifies a very small positive value, in this case 0.05. I_d denotes the identity matrix for the space size, d. Note that, computation of $\sum_i$ when the iteration becomes large turns to be very time consuming and adversely kills the process. Therefore, instead of computing the empirical covariance matrix at iteration i, the online empirical covariance, $\sum_{(i,online)}$, is recommended.

$$\sum_{(i,onlin)} = (\sum_{(i-1,online)}(i-1) + (\frac{i-1}{i})\,(\theta_{i,j} - \mu_{(i-1,j,online)})(\theta_{i,j'} - \mu_{(i-1,j',online)}))/i \quad (3.8)$$

Where, $\mu_{(i,j,online)}$ obtains from Eq.(3.9).

$$\mu_{(i,j,online)} = (\mu_{(i-1,j,online)}(i-1) + \theta_{i,j})/i \quad (3.9)$$

To implement AM the following steps are outlined.

- Initiate the random variables' starting values
- Draw candidate sample exerting Eq.(3.7), Eq.(3.8) and Eq.(3.9)
- Accept or reject the candidate sample according to the MH criterion
- Store the accepted sample
- Proceed to the next iteration

3.3.4. Scaled Adaptive Metropolis (SAM)

Haario et al. (2001),; Andrieu & Thoms (2008); and Roberts & Rosenthal (2009) separately discussed the application of scaling factor, γ_i, in AM. The use of scale factor allows to vary the proposal tune factor (step size) by applying a very small constant value, ε, to approach to the optimum acceptance rate. In this study, after several trials, ε sets to 1e-2 for a fix number of iterations (batch size=500). Eq.(2.19) delivers the condition required to drive the scale factor. Note that, the same comment which is expressed in the PSAMH scale factor is also valid here.

$$\gamma_i = \begin{cases} \gamma_{i-1} + \varepsilon & \xi_i > \xi_{target} \\ \gamma_{i-1} - \varepsilon & \xi_i < \xi_{target} \end{cases} \tag{3.10}$$

Moreover, applying Eq.(2.19), it becomes plausible by redefining Eq.(3.7) to formulate Eq.(3.11)

$$f(\theta_i|\beta, d, \sum_i) = \begin{cases} N(\theta_{i-1}, (\gamma_i 0.1)^2 I_d/d) & i \leq 2d \\ (1-\beta)N\left(\theta_{i-1}, \dfrac{(\gamma_i 2.38)^2 \sum_i}{d}\right) + \beta N\left(\theta_{i-1}, \dfrac{(\gamma_i 0.1)^2 I_d}{d}\right) & i < 2d \end{cases} \tag{3.11}$$

Method implementation procedure described as follows

- Set starting quantities of sampler
- Draw the candidate sample from Eq.(3.11)
- Accept the candidate sample exploiting the MH criterion
- Assess the acceptance rate at batch size and update the scaling factor γ_i, Eq. (2.19), by appreciating the optimum acceptance rate, ξ_{target}
- Store the accepted sample and proceed to the next iteration

3.3.5. Parallel Tempering (PT)

Parallel Tempering (PT) algorithm captures the target distribution by constituting two strata acceptance-rejection conditions. The first criterion allows generating m parallel samples, while the second condition provides a replacement term which is often invoked as the swapping move. In this study, standard deviations of target distribution construct m parallel tempers and they are defined such that always the required standard deviation locates at the center of predefined vector. The diversity of standard deviation vector is illustrated in Table 3.3. In order to implement the PT method, following steps are proposed (Hukushima and Nemoto 1996; Wang and Swendsen 2004; Earl and Deem 2005).

- Set the initial quantities of random field
- Identify the number of m parallel chains and mean of target distribution, μ, which in this case assigned by the value of one. Also, generate the tempering vector which in this case is the standard deviation of target distribution
- Draw candidate samples from the Multivariate Normal distribution supplying the identify matrix as the proposal distribution covariance for m parallel chains (Eq.(2.6) and Eq.(3.12)).

$$\sum = \tau I_d \tag{3.12}$$

Note that, in this study, after several trial and errors, the tune factor (step size), τ, of the proposal distribution takes the value of 2 to enable the sampler to readily explore the entire parameter space.

- Investigate the state of generated samples with the MH criterion (Eq.(3.13)) and store the accepted samples

$$\alpha_{a,PT} = \min(1, \frac{\pi(\theta_{i,k}|\mu, \sigma_k)}{\pi(\theta_{i-1,k}|\mu, \sigma_k)}) \quad (3.13)$$

Where, i and k are the current iteration and number of parallel chains, respectively. μ and σ_k, denote the expected value and the vector of standard deviation of the proposal distribution, respectively, defined in Table 3.3. Note that σ_k indicates a vector of tuning quantities that deterministically aims to alter the target distribution.

- Set the swapping move by drawing a parallel chain number uniformly from the vector of m number of parallel chains, and exclude the target chain number. Bear in mind that, the target chain number indicates the chain that sustains the desired target standard deviation
- Generate the following probability condition and accept it when $\alpha_{b,PT}$ is larger than uniform random value

For each pair of concurrent chain, k and k', which are randomly selected from the swap move

$$\alpha_{b,PT} = \min(1, \frac{\pi(\theta_{i,k}|\mu, \sigma_k) * \pi(\theta_{i,k'}|\mu, \sigma_{k'})}{\pi(\theta_{i,k}|\mu, \sigma_{k'}) * \pi(\theta_{i,k'}|\mu, \sigma_k)}) \quad (3.14)$$

In this case, k denotes the target chain and k' indicates the uniformly selected integer number from the remaining concurrent chains.

- If the above condition is accepted, substitute $\theta_{i,k}$ with $\theta_{i,k'}$.
- The above move iterates to the last chain

- Store the final $\theta_{i,k}$ as the current accepted sample
- Proceed to the next iteration.

3.3.6. Multiple Tries Metropolis (MTM)

In analogy to PT, the MTM technique (Liu, Liang, and Wong 2000; Craiu and Lemieux 2007) also employs several synchronous chains to alter the target distribution. However, the mechanism required to vary the target distribution induced through the weighted samples. Notice that, in contrast to other samplers, however, it is observed that the generated candidate samples varied in the broaden range which justified applying some constraint on the proposal distribution; hence, the proposal distribution inevitably becomes asymmetrical. In order to implement the MTM algorithm, succeeding steps are delineated.

- Set the starting random variables quantities
- Draw m candidate samples from the Normal distribution
- Obtain the value of weight, ω given Eq.(3.15)

$$\omega\left(\boldsymbol{\theta}_{i,k}\right) = \pi(\boldsymbol{\theta}_{i-1}) f(\boldsymbol{\theta}_{i,k}|\boldsymbol{\theta}_{i-1}) \lambda(\boldsymbol{\theta}_{i,k}, \boldsymbol{\theta}_{i-1}) \tag{3.15}$$

Where, $\pi\left(\boldsymbol{\theta}_{\ldots}\right)$ and $f(.|.)$ denote the target and proposal distribution functions, respectively. $\lambda(,)$ indicates a nonnegative symmetrical function which in this case is achieved by Eq.(3.16).

$$\lambda(\boldsymbol{\theta}_{i,k}, \boldsymbol{\theta}_{i-1}) = \left(f\left(\boldsymbol{\theta}_{i,k}\middle|\boldsymbol{\theta}_{i-1}\right) f\left(\boldsymbol{\theta}_{i-1}\middle|\boldsymbol{\theta}_{i,k}\right)\right)^{-\vartheta} \tag{3.16}$$

When, ϑ accepts the value of 1, $\omega(\boldsymbol{\theta}_{i,k})$ turns to normalized target distribution defined by Eq.(3.17).

$$\omega(\boldsymbol{\theta}_{i,k}) = \pi(\boldsymbol{\theta}_{i-1}) f(\boldsymbol{\theta}_{i-1} | \boldsymbol{\theta}_{i,k})^{-1} \tag{3.17}$$

Bear in mind that other alternatives is also presumable for $\lambda(\boldsymbol{\theta}_{i,k}, \boldsymbol{\theta}_{i-1})$, however in this case the recommended combination by Liu et al. (2000) is selected.

- Set $\boldsymbol{\theta}_{i,candidate} = \boldsymbol{\theta}_{i,k}$, such that $\boldsymbol{\theta}_{i,k}$ is drawn randomly from $\boldsymbol{\theta}_{i,\cdot} = \{\theta_{i,1}, \dots, \theta_{i,m}\}$ proportional to $\omega(\boldsymbol{\theta}_{i,k})$
- Draw $m-1$ candidate samples $\{\hat{\theta}_1, \dots, \hat{\theta}_{m-1}\}$ from the proposal distribution with $\mu = \theta_{i,candidate}$
- Set $\hat{\theta}_m = \theta_{i-1}$ as the m^{th} candidate sample
- Accept $\theta_{i,candidate}$ with the following probability

$$\alpha_{MTM} = \min\left(1, \frac{\sum_{k=1}^{m} \omega(\theta_{i,k})}{\sum_{k=1}^{m} \omega(\hat{\theta}_k)}\right) \tag{3.18}$$

- Proceed to the next iteration

Following, the techniques required to evaluate the performance of samplers are outlined.

3.4. Techniques to assess the performance of samplers

3.4.1. Integrated autocorrelation time function

Integrated autocorrelation time function (ACT) (Thompson 2010a) permits to compare the efficiency of a sampler, which is specified via Eq. (3.19).

$$\tau = 1 + 2\sum_{\delta=1}^{\infty} \rho_f(\delta) \tag{3.19}$$

Where, $\rho_f(\delta)$ denotes the correlation coefficient (Box, Jenkins, and Reinsel 2008) and produces a number between -1 and 1and is calculated using Eq.(3.20).

$$\rho_f(\delta) = \frac{\sum_{i=1}^{n-\delta}(x_i - \mu)(x_{i+\delta} - \mu)}{\sum_{i=1}^{n}(x_i - \mu)^2} \tag{3.20}$$

n indicates the number of iterations while δ presents the lag between two randomly generated samples.

Thompson (2010b) described several approaches to attain ACT. In this experiment, the method of initial positive sequences (IPS) is selected which truncates the sum of the autocorrelation coefficient by removing the sum of adjacent negative quantities. The smaller IPS denotes the more efficient sampler in the sense of generating less dependent random samples, and identified with the number of effective samples, $n_{eff.}$.

$$n_{eff.} = \frac{n_{Burn-in}}{\delta_{ACF=0}} \tag{3.21}$$

Where, $n_{Burn-in}$ presents the number of generated random samples after stationary (Burn-in point).

3.4.2. Number of lags ACF=0

Furthermore, the number of lags (δ) which defines a condition that the autocorrelation function (ACF) becomes zero, is computed and shown as an extra comparison criterion. The lag corresponding to the ACF=0 expresses the state of independency of random

variables and is an essential component for computing the number of effective samples, n_{eff}. Similar to ACT, the smaller lag denotes a better performance.

3.4.3. CPU running time

Another possibility to assess the efficiency of sampler appears as the CPU running time that implicitly indicates the effort should be taken to run the MCMC sampler. Smaller quantity of CPU running time, hence, exhibits a better sampler.

3.4.4. Modes recognition based on the relative frequency histogram

The next alternative is the ability of a sampler in properly capturing all modes of target distribution when the standard deviation and correlation coefficients amidst random field varied. Therefore, the plot of relative frequency histogram can be employed to indicate the capability of sampler regarding capturing all target distribution modes.

3.4.5. CDF plot of experimental and true distribution

The last option to feature the performance of samplers is the comparison plot of the cumulative density function of both the random samples and the genuine target distribution.

3.5. Experimental design

A synthetic case constitutes a tetra-modal target distribution in order to elucidate the performance of aforementioned frameworks. A tetra-modal target distribution incorporating the univariate (θ) and bivariate (θ_1, θ_2) random field serves to delineate the characteristics of the target distribution. Table 3.1 categorizes the experimental cases regarding the tetra-modal target distribution case.

Table 3.1 Experimental design of univariate and bivariate target distribution

Univariate ($\boldsymbol{\theta}$)		Bivariate (θ_1, θ_2)		
Case	Standard deviation, σ	Case	Standard deviation, σ	Correlation Coefficient, ρ
$CU1$	1	$CB1$	1	0
$CU2$	0.5	$CB2$	0.5	0.3
$CU3$	0.1	$CB3$	0.1	0.8

The use of various standard deviation and correlation coefficient permits to constitute the case study.

3.6. Univariate and bivariate tetra-modal target distributions

A tetra-modal target distribution comprising one (θ) and two (θ_1, θ_2) random parameters designed to address the performance of PSAMH framework together with several MCMC well-known samplers. The tetra-modal target distribution is formulated in Eq.(3.22).

$$\pi(\boldsymbol{\theta}) = \sum_{e=1}^{4} \omega_e * N(\boldsymbol{\theta}, \mu_{\boldsymbol{e}}, \sigma) \tag{3.22}$$

Eq. (3.22) projects a tetra-modal distribution with a constant standard deviation, σ and various mean, μ_e. $\boldsymbol{\theta}$ denotes a vector of random variables constrained, in this case, in the range of -10 to 10. ω_e provides a vector of magnifying coefficients to construct the height of modes. The associated constant quantities which are common in all samplers are illustrated in Table 3.2.

Table 3.2 Input values of tetra-modal target distribution

Parameter	Value
μ	-6, -2, 2, 6
σ	1, 0.5, 0.1
ρ (Bivariate)	0, 0.3, 0.8
ω	0.2, 0.4, 0.1, 0.3

σ and ρ denote two key altering components that begin with an easy to capture ($\sigma = 1$ and $\rho = 0$) to extremely challenging case ($\sigma = 0.1$ and $\rho = 0.8$). Applying the last case, allows to generate a target distribution with distanced modes, and that is for, it becomes difficult for most samplers to explore the entire random field and hence consequently stick in one or two modes.

Additionally, according to the definition, it is known that the standard deviation supervises the dispersion of a distribution. A larger standard deviation, therefore, constitutes a wider distribution and eventually, in the case of constraint multimodal distribution, modes become connected together; whilst a smaller standard deviation concludes in a slimmer distribution. The distance between modes, on the other hand, can be interpreted as the level of energy or temperature of distribution. It is evident that further distance between modes implies the higher level of energy and is more difficult for sampler to jump from one mode to another one. Provided that, Fig. 3.1 demonstrates the configuration of the univariate experiment for standard deviations accepts values of 0.1, 0.5, and 1 which states the dissemination of distribution modes when the standard deviation varies.

Applying the same mechanism, the significance between the univariate and bivariate experiments appears in inducing the correlation structure amidst the random variables in the bivariate target distribution.

Fig. 3.2 aims to delineate the influence of variations in 3D and 2D bivariate target distributions and to visualize the state of modes when the standard deviation (σ) and correlation coefficient (ρ) are manipulated.

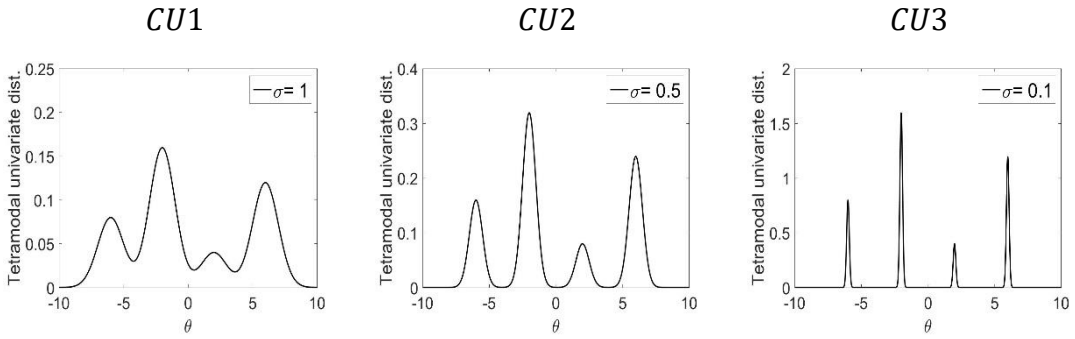

Fig. 3.1. The configuration of univariate tetra-modal target distribution with various standard deviations

Additionally, the projection of 3D distributions on the side planes demonstrates the histograms of random variables (Fig. 3.2, left). Also, considering the 2D joint histogram in Fig. 3.2, right, it is evident that by decreasing the standard deviation and increasing the correlation coefficient gaps between modes are expanded, which consequently make it more difficult for the sampler to jump from one mode to another.

3.7. Samplers input values

Several assigned initial quantities are earlier provided when each sampler is elaborated. Nevertheless, the remaining necessary input data to be able to implement samplers are identified subsequently.

3.7.1. PSAMH

The number of concurrent chains, m, in this method is set to 5 and 10 for univariate and bivariate target distribution cases, respectively. The initial searching values are selected randomly but investigated that they provide a non-zero values when they are inserted in the tetra-modal distribution. It is known that the adaptive methods would trap in a point without progress, if wrong initial searching values are picked.

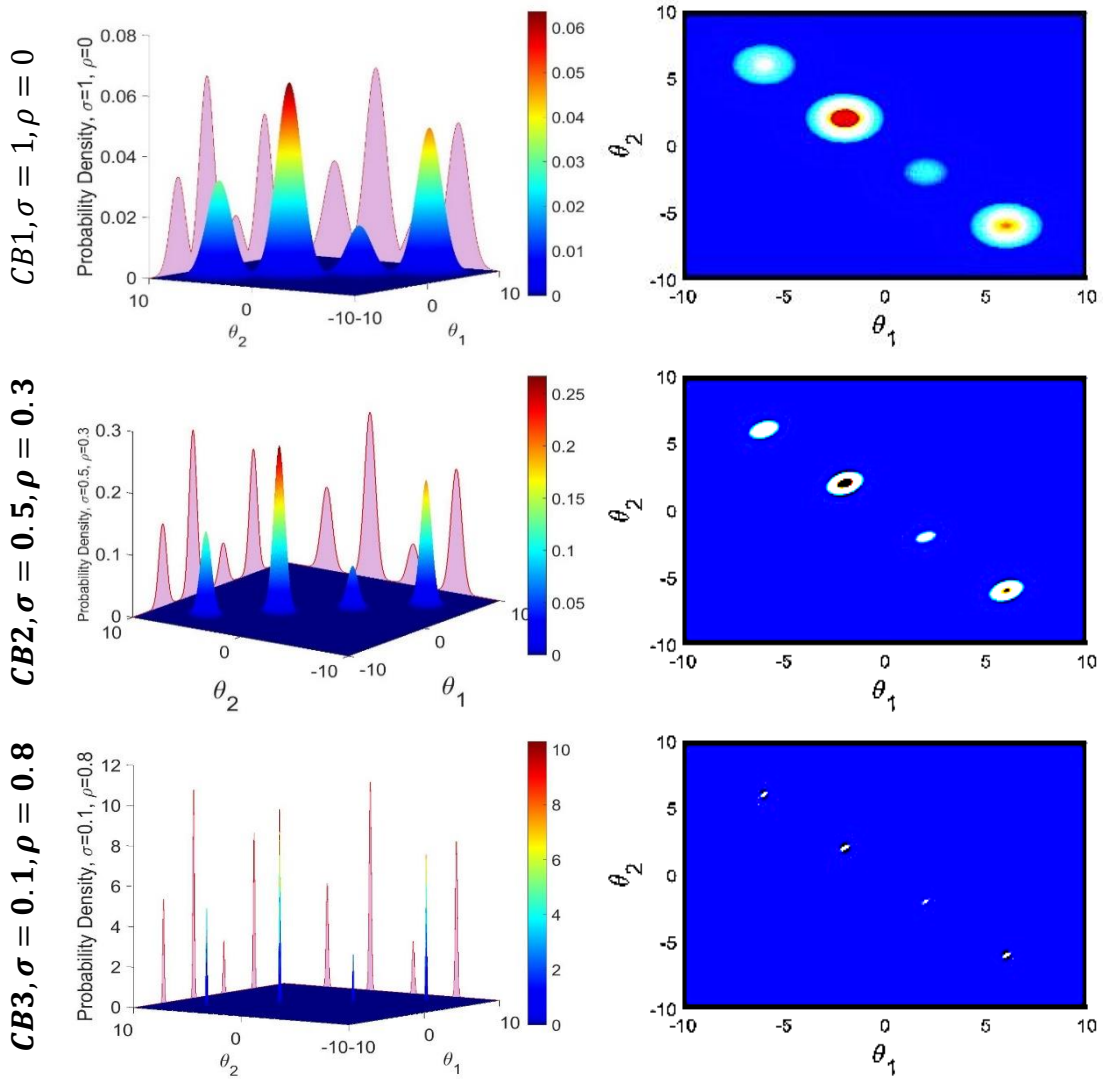

Fig. 3.2. Left, 3D and right, 2D configuration of bivariate tetra-modal distribution

3.7.2. PT

Table 3.3 illustrates the input quantities of PT method.

Table 3.3 PT input data

σ	Univariate, $m = 5$	Bivariate, $m = 9$
0.1	0.05, 0.075, 0.1, 0.2, 0.3	0.02, 0.04, 0.06, 0.08, 0.1, 0.12, 0.14, 0.16, 0.18
0.5	0.1, 0.3, 0.5, 0.7, 0.9	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
1	0.4, 0.7, 1, 1.3, 1.6	0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8

In each of above cases, temperatures are specified such that they allow the sampler to readily explore the parameter space providing a wide range of constraints.

3.7.3. MTM

The tune factor values and number of tries corresponding to each case study is presented in Table 3.4.

Table 3.4 MTM input values

	Univariate	Bivariate
m	10	10
Tune factor	3	2

Defining the entire initiate values and conditions permit to drive MCMC samplers. The associated results are provided and discussed in the next section.

3.8. Results and discussion

Following tables present the results obtained from the implementation of aforementioned sampling methods. The tables constructed such that, the name of samplers is indicated in the column one. Moreover, the ability of samplers to capture all tetra modes is denoted in the second column. Columns three and four illustrate the acceptance rate and CPU running time, respectively. Column five provides the number of lag when the auto correlation function takes the value of zero. Finally, the last column shows the retrieved value of IPS.

The subsequent figures demonstrate some plots in the proceeding order. The first column indicates the name of each sampler. The MCMC experiment plots are depicted in the second column. In addition, the plots of relative frequency histogram and the

comparison between the cumulative density function of empirical (gray dashed line) and true target distribution (black solid line) are exhibited in the last two columns.

3.8.1.1. $CU1$: $\sigma = 1$

$CU1$ denotes the univariate tetra-modal distribution when the standard deviation sets as 1(Table 3.5). In this case, all samplers enabled to successfully capture posterior modes (Fig. 3.3). RWM required less running time, which is the basic sampler. However, AM and SAM together presented the smallest values of lags corresponding to ACF=0, while PT provided the smallest ACT. PSAMH and SAM developed acceptance rates close to the optimum acceptance, which in this case is %44.2.

Table 3.5 Results of case $CU1$, σ = 1, univariate

Method	Able to Capture all modes	Acceptance rate (%)	Running Time (s)	ACF=0	ACT (IPS)
PSAMH	Yes	43.13	62.28	80	556.71
PT	Yes	64.57	257.37	80	296.1
MTM	Yes	89.21	98.18	20	509.75
AM	Yes	34.32	53.07	10	381.98
SAM	Yes	44.17	50.87	10	431.16
RWM	Yes	73.8	30.39	230	1734.8

The point that is worthwhile to mention in the MCMC experiment (Fig. 3.3) regarding the PSAMH technique is the appeared gaps in random samples. The gaps indicate those iterations that the sampler scaled the step size to approach to the optimum acceptance rate. In other words, for several intervals the sampler rejected more random samples than other iterations to achieve the optimum acceptance rate. In addition, other results inferenced from Fig. 3.3 summarized as the relative frequency histogram plots display the successful performance of samplers in capturing all posterior modes.

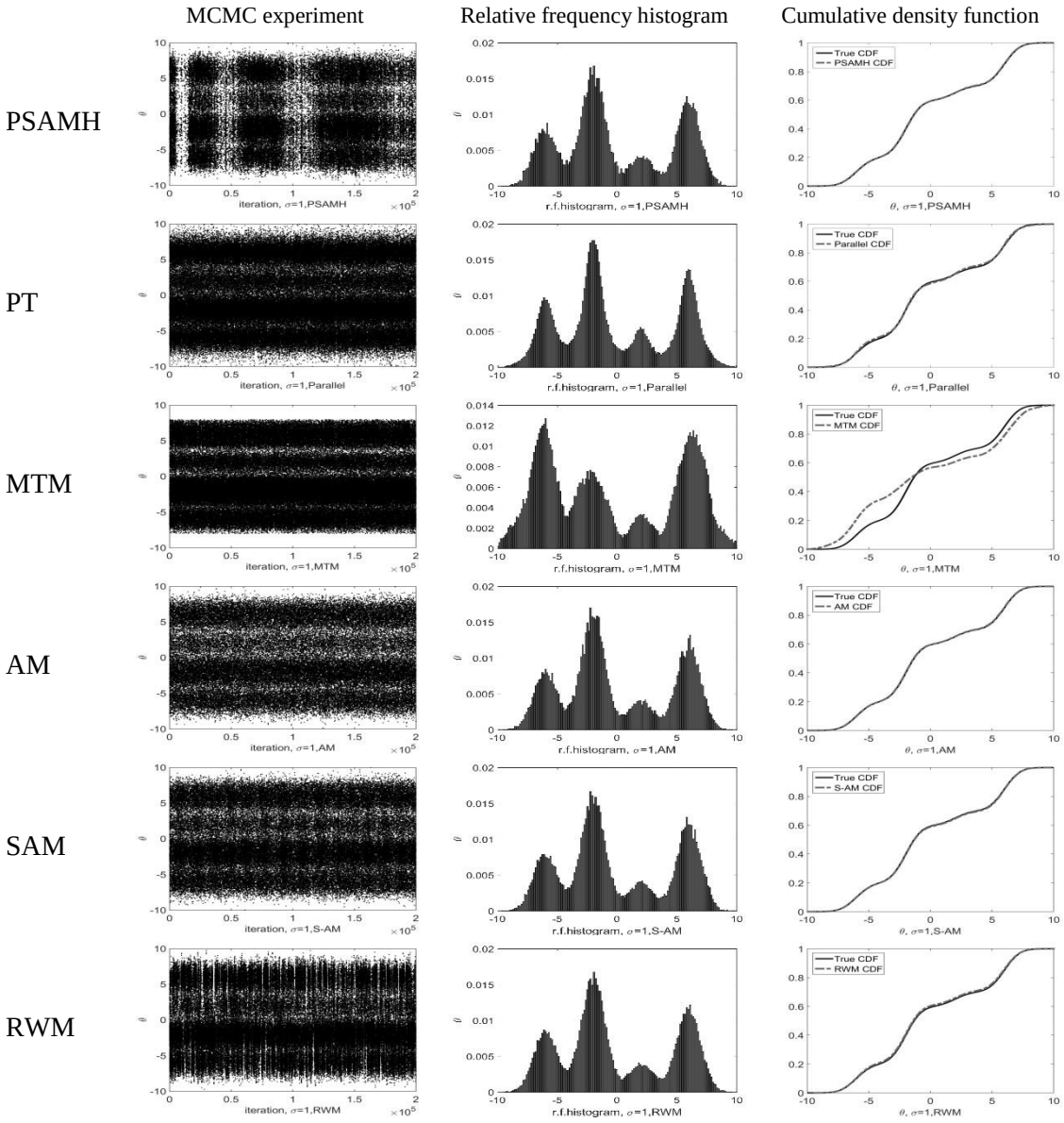

Fig. 3.3. Experiment and statistics results for $CU1$

Furthermore, cumulative density function plots demonstrate an acceptable agreement between the empirical CDF and target distribution CDF. Nevertheless, CDF of MTM method exhibits some degree of distortion in compare to the target CDF.

3.8.1.2. $CU2$*:* $\sigma = 0.5$

Computation results in Table 3.6 reveals the capability of all samplers to independently explore the posterior modes when σ = 0.5. RWM provided less running time but the highest lag to ACF=0 and ACT. PT demanded more running time, however, attained ACT became slightly smaller than PSAMH. MTM presented less sensitivity to the variation of standard deviation since the acceptance rate remained unchanged. Nevertheless, except PSAMH, SAM and MTM other samplers experienced more rejection in compare to $CU1$.

Table 3.6 Results of case $CU2$, σ = 0.5, univariate

Method	Able to Capture all modes	Acceptance rate (%)	Running Time (s)	ACF=0	ACT (IPS)
PSAMH	Yes	45.78	70.71	10	379.98
PT	Yes	36.9	284.4	260	359.7
MTM	Yes	80.18	91.43	10	508.25
AM	Yes	18.13	47.3	30	548.34
SAM	Yes	44.18	51.41	1910	2306
RWM	Yes	50.59	30.04	2310	5413.7

According to Fig. 3.4, MTM and SAM methods illustrate more deviation regarding the CDF plots than other samplers. Despite MTM is capable to inspect all target distribution modes, the trend of CDF plots not entirely coincident with each other.

3.8.1.3. $CU3$*:* $\sigma = 0.1$

Table 3.7 coupled with Fig. 3.5 evidenced that AM, SAM, and RWM disabled to capture all posterior modes when σ became extremely small. Despite the optimized acceptance rate of SAM, this sampler still stuck in one mode.

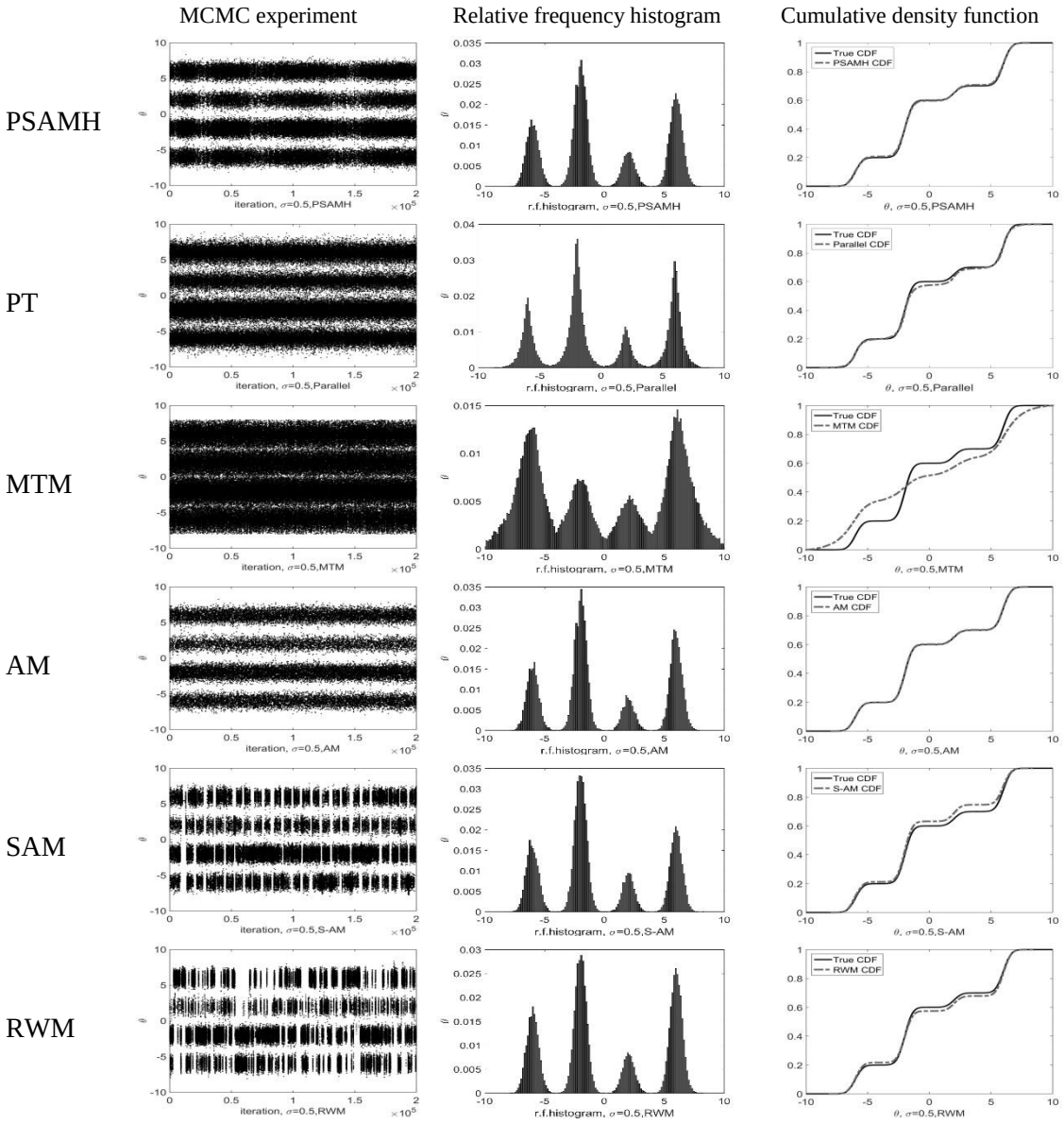

Fig. 3.4. Experiment results for $CU2$

Therefore, approaching to an optimized acceptance rate solely is not able to assure the properly exploring of the entire posterior space to capture all modes. PSAMH and PT, in general, produced acceptable results in compare to other samplers.

Table 3.7 Results of case $CU3$, σ = 0.1, univariate

Method	Able to Capture all modes	Acceptance rate (%)	Running Time (s)	ACF=0	ACT (IPS)
PSAMH	Yes	43.13	62.28	80	556.71
PT	Yes	10.1	294.2	580	633.4
MTM	Yes	57.83	96.81	40	599.9
AM	No	45.47	142.65	-	-
SAM	No	44.2	137.1	-	-
RWM	No	24.46	31.9	-	-

In the case of $\sigma = 0.1$, as seen in Fig. 3.5, AM, SAM, and RWM failed to capture all posterior modes. MTM also suffers from the lack of proper jumps through the parameter space. PSAMH and PT methods displayed appropriate jumps and acceptable adaption to the CDF plots.

3.8.1.4. $CB1$: $\sigma = 1, \rho = 0$

The same procedure that applied to the univariate case is also exerted regarding the bivariate tetra-modal target distribution with altering the correlation structure between random variables in addition to the standard deviation. Table 3.8 introduced the numerical computational results of the case $CB1$.

Table 3.8 Results of case $CB1$, σ = 1 and $\rho = 0$, bivariate

Method	Able to Capture all modes	Acceptance rate (%)	Running Time (s)	ACF=0		ACT (IPS)	
				θ_1	θ_2	θ_1	θ_2
PSAMH	Yes	34.51	259.9	150	150	944.9	942.2
PT	Yes	33.75	2237.2	280	280	306.7	303.6
MTM	Yes	51.9	2095.3	440	1000	1876.2	1855.3
AM	Yes	21.9	263.2	80	70	556.1	558.8
SAM	Yes	34.57	266.5	1450	1380	1061.5	1134.8
RWM	Yes	49.1	182.2	1410	1420	4178.8	4173.8

Considering Table 3.8, it is evident that all samplers were able to properly capture all posterior modes.

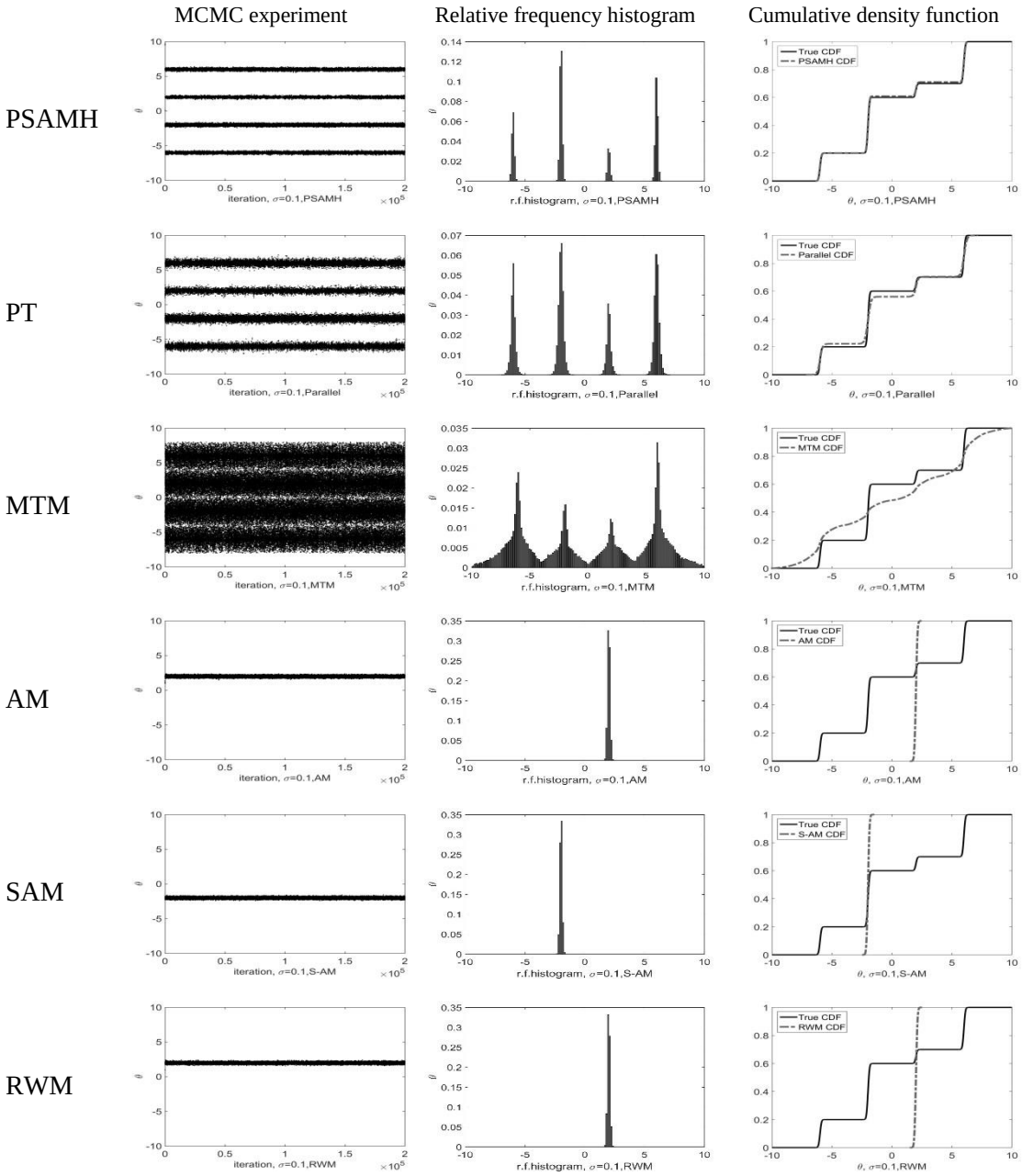

Fig. 3.5. Experiment results obtained from $CU3$

Although, RWM exhibited less running time, PSAMH, AM and SAM took the next place. As it is expected, both PSAMH and SAM delivered acceptance rate close to the optimum. Although, PT produced smaller values of IPS but suffered from the noticeably

longer running time. In addition, AM and then PSAMH provided more independent samples in comparison to the other samplers. The quantities of IPS and lags of ACF=0 are varied depend on the samplers, however, RWM produced larger values.

The relative frequency histogram of all methods in Fig. 3.6 except MTM demonstrated an acceptable behavior. Even though, MTM captured all modes, the state of modes is not distinguishable. The same inference can also be implied by observing the CDF plots.

3.8.1.5. $CB2$: $\sigma = 0.5, \rho = 0.3$

Table 3.9 demonstrates the outcomes of the moderate case where standard deviation and correlation coefficient get the values of 0.5 and 0.3, respectively. Considering the ability of samplers in capturing all modes, it is revealed that, RWM, AM and SAM disabled to appropriately explore the entire random space.

Table 3.9 Results of case $CB2$ **,** σ = 0.5 and $\rho = 0.3$, bivariate

Method	Able to Capture all modes	Acceptance rate (%)	Running Time (s)	ACF=0		ACT (IPS)	
				θ_1	θ_2	θ_1	θ_2
PSAMH	Yes	33.84	257.77	360	350	1466.8	1470.7
PT	Yes	11.76	2261.9	1360	1360	550.6	555.3
MTM	Yes	40.02	2030.4	790	800	2292.9	2305.6
AM	No	37.22	268.8	-	-	-	-
SAM	No	34.8	275.6	-	-	-	-
RWM	No	40.9	170.5	-	-	-	-

The running time of samplers outweighed the technique of PSAMH in compare to PT and MTM. PSAMH additionally provided ACF=0 with smaller magnitudes, while PT illustrated better IPS. In general, the use of PSAMH method allowed to both save the running time and generate more independent samples.

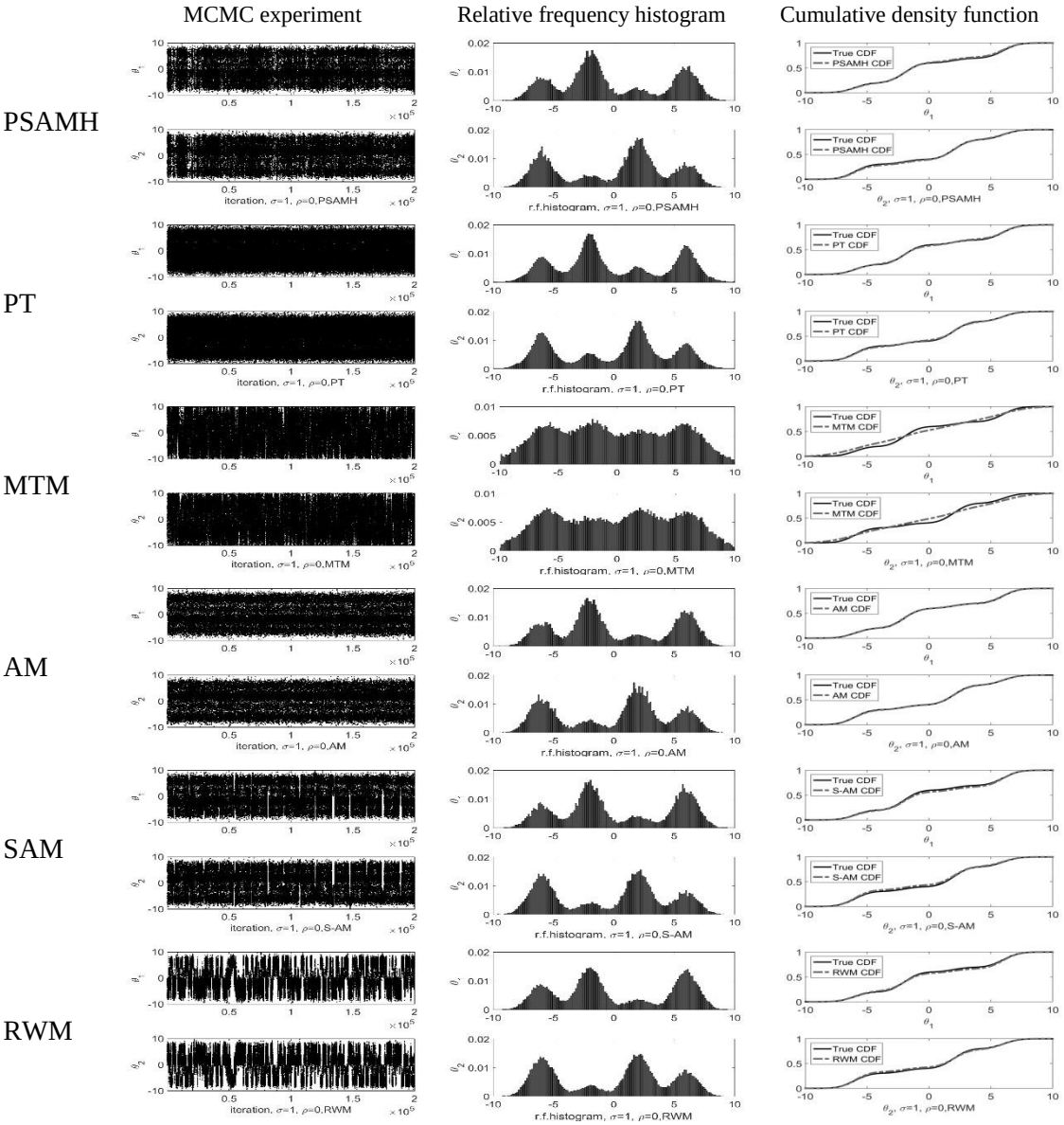

Fig. 3.6. Experiment results for $CB1$

A new stage of difficulty is depicted in Fig. 3.7 as AM, SAM, and RWM grinded to a halt in capturing all modes. Furthermore, the relative frequency histogram plots demonstrated the uncertain in modes captured by MTM algorithm. Regardless of the

degree of difficulty presented in this case, either, the CDF plots or relative frequency histograms depicted the capability of PSAMH and PT to readily jump in the parameter space and capture all target modes.

3.8.1.6. $CB3$*:* $\sigma = 0.1, \rho = 0.8$

By progressively increasing the correlation coefficient between random variables, the running time, lags of ACF=0 and IPS escalated dramatically in all samplers (Table 3.10). In this case, AM, SAM and RWM, in analogy to the previous case, were not able to explore the posterior space properly. PSAMH, however, generated random samples in a more efficient computational time. Nevertheless, MTM provided substantially smaller values of ACF=0. Also, while PSAMH sustained the acceptance rate close to the optimum quantity, PT surprisingly rejected most of generated samples.

Table 3.10 Results of case $CB3$ **,** σ = 0.1 and $\rho = 0.8$, bivariate

Method	Able to Capture all modes	Acceptance rate (%)	Running Time (s)	ACF=0		ACT (IPS)	
				θ_1	θ_2	θ_1	θ_2
PSAMH	Yes	32.62	257.8	41880	41880	23887	23882
PT	Yes	0.5	2259.4	20240	20240	4420.5	4444
MTM	Yes	61.32	1909.3	510	910	2317.7	2214.5
AM	No	28.3	300.7	-	-	-	-
SAM	No	35.03	250.8	-	-	-	-
RWM	No	16.03	175.14	-	-	-	-

In contrary, MTM stated more acceptance rate when it compared with the previous case. PT and PSAMH presented extremely large numbers regarding the ACF=0 and IPS.

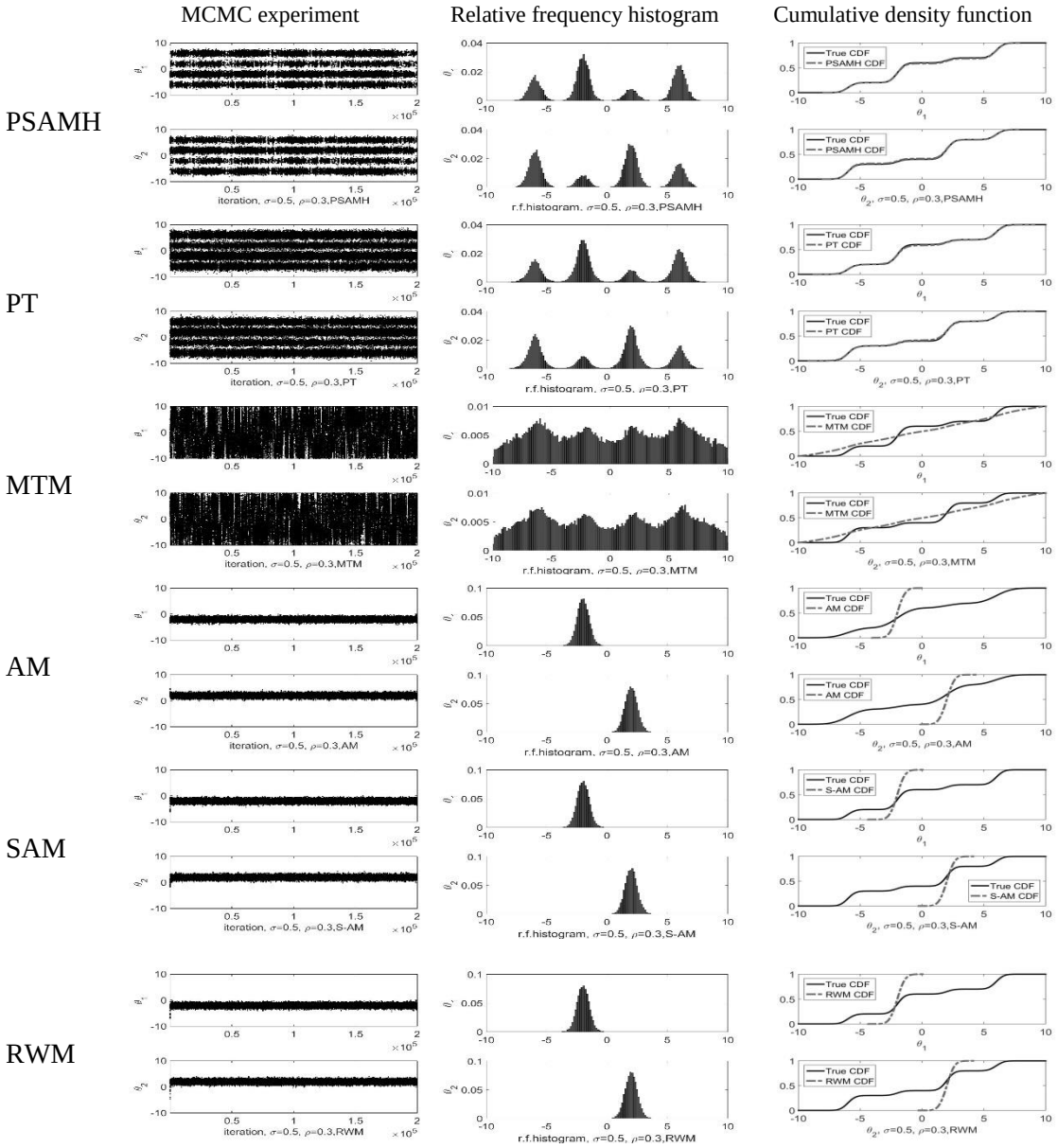

Fig. 3.7. Experiment results for $CB2$

Fig. 3.8, additionally, appreciated the results expressed in Table 3.10. MTM presented persistency in the ill jump pattern by removing the gap between target distribution modes. PSAMH and PT yet exhibited successful jumps in the posterior space. However,

considering the trend of CDF plots, PSAMH sampled the entire space with more precision than PT.

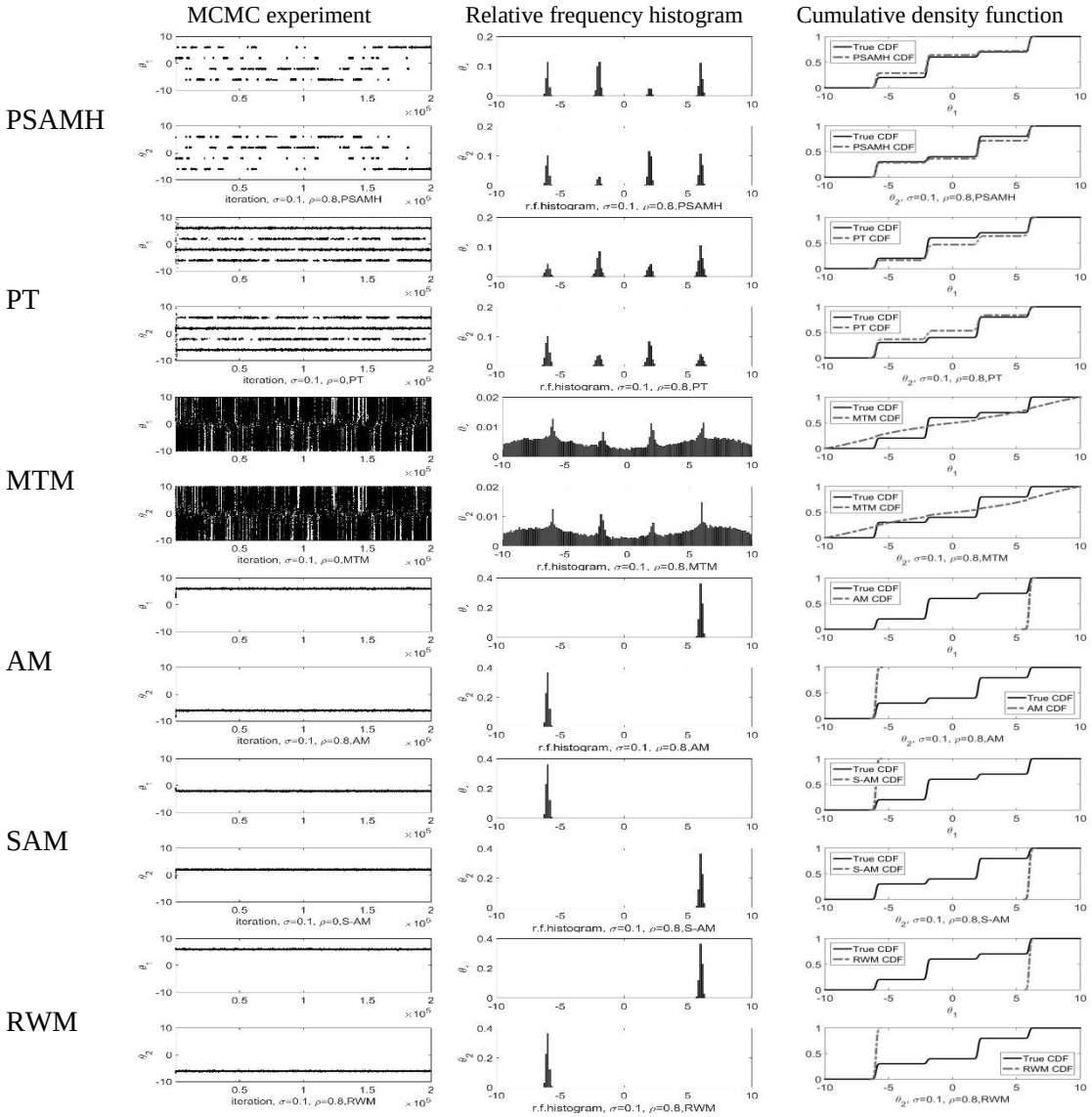

Fig. 3.8. Experiment results for $CB3$

3.8.2. Acceptance rate and scaling factor in PSAMH and SAM

Graphical comparisons of the case of univariate target experiment amidst the PSAMH and SAM techniques in the acceptance rate and scale factor are demonstrated in Fig. 3.9.

SAM modified the acceptance rate relatively faster while depicting more fluctuation regarding the scale factor. The scale factor of PSAMH, additionally, provides a numeric to implicitly assess the state of stationary condition for chains where after some iteration, it becomes stable around number one. The noticeable oscillation of scale factor in SAM for $\sigma = 0.1$, clearly depicts the level of difficulty imposed to the sampler. A brief conclusion is drawn in section 8.1.2.

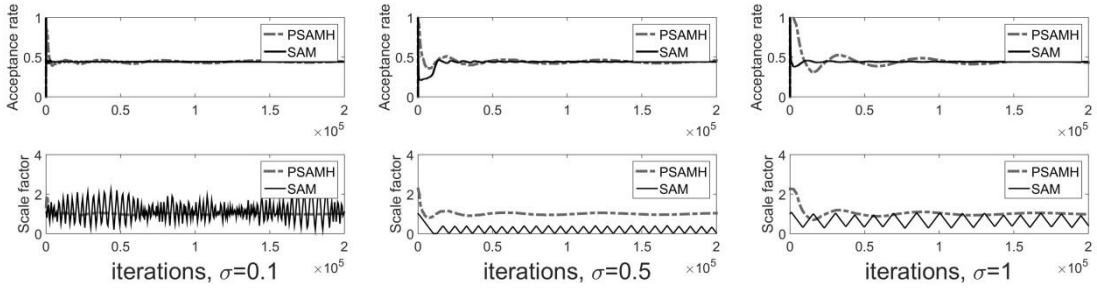

Fig. 3.9. Comparison of acceptance rate and scale factor between PSAMH and SAM

4. A NOVEL PROBABILISTIC BAYESIAN APPROACH TO QUANTIFY THE UNCERTAINTY ASSOCIATED WITH THE MODIFIED ARPS MODELS

4.1. Overview

Empirical or Arps-based models allow us to assess the Estimated Ultimate Recovery (EUR) of wells by matching the likely model approximation to the completion logs, accounting only the general trend of productions. Despite their wide application due to the simplicity in the implementation, scholars yet endeavor to extent a practical technique that precisely quantifies the inherent uncertainty in the EUR analysis when the statistical comprehension of the incorporated model parameters appears to be obscure. Hence, that is for, it is pertinent to study the model performance via exerting the Bayesian paradigm. The Bayesian probabilistic analysis coupled with the Markov Chain Monte Carlo (MCMC) and Metropolis-Hastings (MH) develops a mean to not only calibrate the model variables but also evaluate the associated uncertainty of well productions. However, it is connoted that often, due to the complexity of the problem, MH becomes deficient of readily sampling the random filed. In such cases, a more sophisticated sampling method such as the Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) framework is required. PSAMH constitutes several concurrent chains to optimize the acceptance rate by adapting the proposal distribution of MH algorithm. Also, to be able to delineate the well's behavior over the course of varying time, two Arps-based models, the Modified

Hyperbolic Decline (MHD) and the Power Law Exponential Decline (PLED) in association with four sets of wells' production data retrieved from the Eagle Ford Shale are exploited. In addition, the vector of posterior acquired from MCMC, permits the extrapolation of model realizations for the short and long term of production. The study eventually implies that, the Bayesian paradigm along with the PSAMH method is enough to provide a metric to quantify the computational uncertainty by ensuring the exploration of the entire parameter space. Moreover, considering the given well depletion data, a comparison amid two modified Arps' models, exhibits the overestimation of EUR together with the confidence's diminution corresponding to the MHD model.

4.2. Introduction

Arps (1944) introduced several empirical decline curve models, which by tuning the model parameters match a decline curvature to well depletion logs to evaluate the estimated ultimate recovery (EUR) of reserves. The simplicity in the implementation of models and noticeably less number of constitutive variables in compare to the analytical models have founded the extended popularity of these models. Despite that Arps' decline, models were initially designed to evaluate the conventional reservoirs productions, later applying some modifications, the same models also exploited for unconventional reservoirs. The modified Arps models exerted in this research are the Modified Hyperbolic (MHD) (Robertson 1988) and Power Law Exponential (PLED) (Ilk et al. 2008) Decline curve models. A comprehensive comparison between two aforementioned models is driven by Seshadri and Mattar (2010). According to the Bayesian literature and hereafter, the empirical models will be called "forward models".

The major difficulty in the EUR assessment methods lays in the quantification of their inherent uncertainty associated with both model parameters and production forecasting (Capen 1976). Recently, several scholars have attempted to address the confidence of modeling's results by employing probabilistic frameworks such as the Bootstrap method (Jochen, Spivey, and Holditch 1996) or Bayesian approaches (Cheng et al., 2010; Xie et al., 2011; Gong et al., 2014; Purvis and Kuzma, 2016). Abdollahzadeh et al. (2011) provided several Bayesian optimization algorithms to quantify the uncertainty associated with applying the history matching method. A Bayesian solution considering a frequent transition between different covariance model in MCMC is introduce by Elsheikh et al. (2012) employing the history matching for oil reservoirs. Moridis et al. (2017) employed the Bayesian paradigm and MCMC to calibrate empirical decline curve models and used the provided statistics to extrapolate the projection for a long-term production.

Bayesian paradigm has become one of the most invoked methods dealing with the uncertainty quantifications (Gong et al. 2014; Vink and Gao 2015). Markov Chain Monte Carlo among several Bayesian approaches has been introduced as a computational framework which by integrating random samples across the entire parameter space and employing some accepting or rejecting criteria such as Metropolis-Hastings (MH) (Metropolis et al. 1953; Hasting 1970) eventually captures the posterior (Gelman et al. 2013). Since it is initially presumed that in most probabilistic analysis problems, it is almost impossible to directly take the sample from the posterior space, MH exploits an auxiliary distribution to as a replacement option generate the random samples from it. The auxiliary distribution is often recalled as "proposal distribution."

The proposal distribution incorporates of two terms, a tuning factor which adjusts the step size of the random movement and the initial random sample (Gareth O. Roberts and Rosenthal 2002).

Adaptive Markov Chain Monte Carlo methodology (Adaptive MCMC) establishes a mechanism which by tuning the step size optimizes the MH acceptance rate (Haario et al., 2001; Rosenthal et al., 2010; Whiteley et al., 2010; Roberts et al., 2012).

Parallel Scaled Adaptive Metropolis Hasting (PSAMH) introduced as a sophisticated framework which augments (see section 2.3.2) the Adaptive MCMC with parallelizing several concurrent chains in order to automatically tune the step size and capture the optimized acceptance rate. In this research, by applying PSAMH, we conducted several experimental cases on the wells depletion logs extracted from the Eagle Ford Shale formation to assess the uncertainty associated with forward models. Furthermore, a comparison between two forward models are derived to generate a better cognition of application of models. The advantage of PSAMH is concisely embedded in the simplicity of implementation and promising features to capture the true posterior of the Bayesian process.

4.3. Methodology

Decline Curve Analysis (DCA) has become a common practical method to simulate the conventional or unconventional plays' depletion for either a short or a long term of well production. According to Arps, there are various decline curve functions, which are suitable for assorted formations. Power Law Exponential (PLED) and Modified

Hyperbolic (MHD) decline curves are among the modified Arps methods, which have recently found their popularity in the case of hydro-fractured reserves.

4.3.1. Modified Hyperbolic Decline (MHD)

Hyperbolic decline curve equation (Fetkovich, Fetkovich, and Fetkovich 1996) is given by

$$q = \frac{q_i}{(1 + bD_i t)^{\frac{1}{b}}} \tag{4.1}$$

Where, q and q_i represent the reserves production rate and the initial rate, respectively. b indicates the exponential (typically between 0 and 1) and D_i denotes the initial decline rate.

In Eq. (4.1), the decline rate, D, is not constant and can be reckoned by Eq. (4.2).

$$D = \frac{1}{\frac{1}{D_i} + bt} \tag{4.2}$$

Note that, when D becomes too small, the gas rate, no longer declines significantly and the reserves can be over predicted. To diminish the mentioned pitfall, D_{limit} is modified and plugged into the Eq. (4.1) as subsequent.

$$q = \begin{cases} \dfrac{q_i}{(1 + bD_i t)^{\frac{1}{b}}} & t \leq t^* \\ q_i \exp(-D_{limit}\, t) & t > t^* \end{cases} \tag{4.3}$$

Eq. (4.3) presents Modified Hyperbolic Decline curve (MHD) (Robertson 1988).

Where,

$$D_{limit} = -\frac{\ln[1 - p]}{365} \tag{4.4}$$

And

$$t^* = \frac{(\frac{D_i}{D_{limit}} - 1)}{bD_i} \quad (4.5)$$

The value of p sets to 10%.

4.3.2. Power Law Exponential Decline (PLED)

Eq. (4.6) defines the Power Law Exponential decline model (Ilk et al. 2008).

$$q = q_i \exp(-D_\infty t - D_i\, t^n) \quad (4.6)$$

Where, n denotes the time exponent (typically between 0 and 1). q_i indicates the initial rate q (t=0). D_∞ and D_1 present the decline rate at infinite time and instantaneous decline at time t=1, (assuming D_∞=0), respectively.

$$D_i = \frac{D_1}{n} \quad (4.7)$$

Typically the value of b in MHD and n in PLED should be between 0 and 1, however, Seshadri and Mattar (2010) denoted that this range is not suitable when the permeability is extremely low and hence it is necessary to increase b or n to more than 1.

In the current research, q_i, b and D_i of MHD and q_i, D_∞, D_i and n of PLED are set as the Bayesian random variables.

4.3.3. Bayesian paradigm and Markov Chain Monte Carlo (MCMC)

The notions of Bayesian paradigm and Markov Chain Monte Carlo along with all equations and functions are similar to what is discussed in section 2.3. Therefore, we avoid reiterating all materials in this section and encourage the reader to study the cited section.

4.3.4. Parallel Scaled Adaptive Metropolis Hasting (PSAMH)

Although, the concept of PSAMH comprehensively is initiated in section 2.3.2 and extra augmentation features are explained in section 2.4 and the implementation sequence is delineated in section 2.5, due to importance of PSAMH framework, a concise computation procedure together with some significant equations are recapitulated in this section.

Authors in section 2.3.2 constructed the PSAMH framework to draw a set of random samples utilizing m concurrent chains considering the optimum acceptance rate, ξ_{target}, of Adaptive MCMC approach. PSAMH elaborates the coefficient of variation (CoV) to regulate the proposal distribution step size. CoV indicates the fraction of the standard deviation (σ) and mean (μ) of generated random variables (Forkman 2009).

$$\mathrm{CoV} = \frac{\sigma}{\mu} \tag{2.12}$$

Subsequent, we briefly describe the required steps should be taken to implement PSAMH, however, we recommend the practitioner to read section 2 to better discern the gist of PSAMH framework prior proceeding to the next section.

- Retrieve the optimum acceptance rate, $\xi_{target}(d)$, from Eq.(2.18).

$$\xi_{target}(d) = 0.234 + \frac{0.654}{(1 + 0.775 * 1.846 * d)^{1/0.775}} \tag{2.18}$$

d indicates the parameter space size or in other words, the number of random variables.

- Draw $[m \times d]$ combination of the coefficient of variation (CoV) from the Chi square distribution with degree of freedom of 1.5 ($Chi2(1.5)$).

- Obtaining the initial values of CoV, construct the covariance matrix, $\sum$ of the Multivariate Normal distribution applying Eq.(2.14) .

$$\sum = \begin{bmatrix} (\mathrm{CoV}_1 * \mu_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathrm{CoV}_d * \mu_d)^2 \end{bmatrix} \tag{2.14}$$

Where, for $i = 1$, μ_j denotes the initial values of forward model parameters whereas for $i > 1$, μ_j is determined by $\boldsymbol{\theta}_{i-1}$.

- At iteration i, assess the state of MH criterion (Eq.(2.3)) in each single m concurrent chains. Determine the final accepted sample among l presumably accepted samples, ($0 \leq l \leq m$), adapting the following condition.

$$\hat{r}_{i,k,k'} = \frac{\pi(\boldsymbol{\theta}_{i,k}|\boldsymbol{d}_{obs})}{\pi(\boldsymbol{\theta}_{i,k'}|\boldsymbol{d}_{obs})} \frac{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k})}{f(\boldsymbol{\theta}_{i,k}|\boldsymbol{\theta}_{i-1})} \frac{f(\boldsymbol{\theta}_{i,k'}|\boldsymbol{\theta}_{i-1})}{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k'})} \tag{2.11}$$

k and k' are two accepted concurrent chains of l total accepted samples. If $\hat{r}_{i,k,k'} > 1$, then $\boldsymbol{\theta}_{i,k}$ would be perpetuated as the accepted sample otherwise, $\boldsymbol{\theta}_{i,k'}$ would be replaced as the current accepted sample. Terms $\frac{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k})}{f(\boldsymbol{\theta}_{i,k}|\boldsymbol{\theta}_{i-1})}$ and $\frac{f(\boldsymbol{\theta}_{i,k'}|\boldsymbol{\theta}_{i-1})}{f(\boldsymbol{\theta}_{i-1}|\boldsymbol{\theta}_{i,k'})}$ are substituted by Eq.(2.17). The above step should be repeated $l - 1$ times to the last accepted concurrent chain.

- Specify the batch size iteration (e.g. 500 or 1000) and evaluate the state of proximity of the current acceptance rate to the optimum acceptance at iterations equal to the batch size by applying Eq.(2.20).

$$\gamma_i = \xi_i / \xi_{target} \tag{2.20}$$

Where, γ_i presents the scaling factor. In addition, impose γ_i via Eq. (2.21) into the covariance of the proposal distribution.

$$\sum = \gamma^2 * \begin{bmatrix} (\mathrm{CoV}_1 * \mu_1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\mathrm{CoV}_d * \mu_d)^2 \end{bmatrix} \tag{2.21}$$

- Other feature appeared in this research is the reduction of redundancy of the number of concurrent chains, m to m' at $i = 10000$. The notion is choosing m' number of more frequent selected of the combination of CoV relying on the histogram of CoV. This feature eventually deals with the computational time efforts (see section 2.4.2).

4.4. Observed data

The observed data in this research retrieved from the Eagle Fort Shale formation and comprises of four wells (Table 4.1). Since, wells produced both oil and gas, the production flow rate unit is set as the Barrels of Oil Equivalent per Day (BOED).

Table 4.1 Well's number and production duration

Well name	Well 1	Well 2	Well 3	Well 4
Production duration (Day)	738	727	721	740

Fig. 4.1 also demonstrates four wells production rate on the daily basis with the normal and log-log scales.

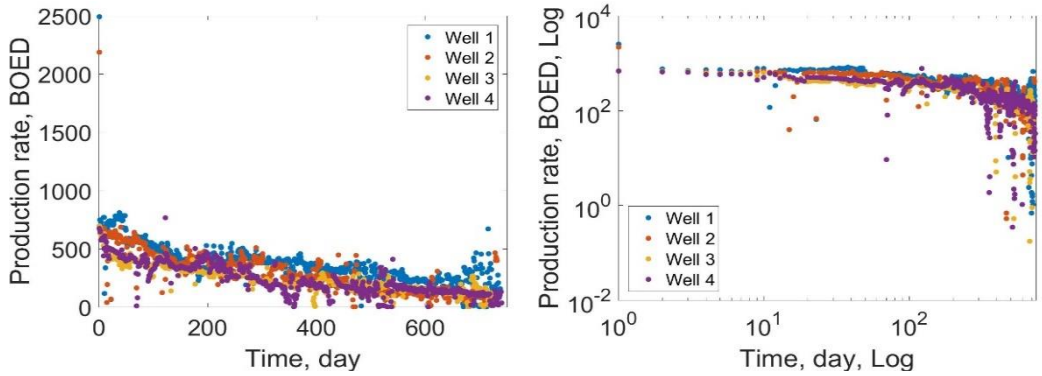

Fig. 4.1. Daily production rate of wells, right, normal scale and left log-log scale

4.4.1. Preparation of observed data

Considering the wells' production plot (Fig. 4.1), some outliers are observed which unnecessarily mislead the general trend of data. In the decline curve methods, moreover, it is recommended to remove outliers to better assess the well behavior. To approximate the possible shape of the error of observed data, which is the tradeoff between the data and fit curve, it is required to compare the empirical cumulative density function (CDF) of all wells with known distribution functions. The Normal and Lognormal distribution functions are served as the error distribution alternatives and the results are depicted in Fig. 4.2, left. The comparison of CDF plots substantiates the meticulous selection of the Normal distribution as the plausible shape of the error. By examining the error of the observed data in proximity of the Nonlinear Least Square (NLS) optimized fit curve, the same conclusion also can be drawn, Fig. 4.2, right. Note that, to evaluate the state of lognormal CDF for error of NLS optimized curve (Fig. 4.2, right), initially the vector of residual is shifted to the positive values - Lognormal function does not accept negative numbers- and after obtaining the CDF, the result is dragged back to the original values.

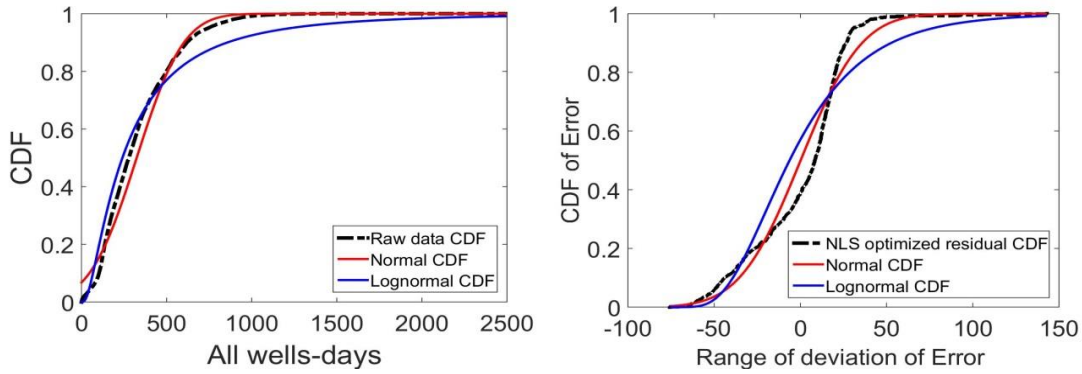

Fig. 4.2. Left, CDF plots of raw data, Normal and Lognormal distribution and right, CDF plots of residual of optimized NLS curve and raw data, Normal and Lognormal distributions

Therefore, for each forward model, a code is scripted to remove the outliers in the several running strata and trim the skewness of data to fit the shape of the error as close as possible to the Normal distribution. The subsequent steps are mapped to filter the raw data and augmented by the following flowchart (Fig. 4.3).

- Set the confidence interval (for example 95%) of the raw data that is presumed to be trimmed, then the error becomes Alpha=5%. The main notion is to remove Alpha percent of raw data along to the skewness.
- Specify the Alpha reduction factor (0.3) to taper the size of Alpha in each loop. This quantity takes a portion of the Alpha value out of the calculation each time. In addition, set the target alpha value (Beta=1e-3). Beta is the minimum error, which is intended to obtain.
- Set the Gamma value (Gamma=3.5). Gamma is used to outweigh on the tail that has the skewness during the filtering.
- Determine the residual of the NLS optimized curve and raw data.

- Identify the skewness of the residual.
- If skewness>0, compute the succeeding conditions

$$\begin{aligned} Lower\ percentile &= Percentile\ (residual, \frac{\text{Alpha}}{2\ Gamma}) \\ Upper\ percentile &= Percentile\ (residual, 100 - \frac{\text{Alpha} * Gamma}{2}) \end{aligned} \tag{4.8}$$

If skewness<0

$$\begin{aligned} Lower\ percentile &= Percentile\ (residual, \frac{\text{Alpha} * Gamma}{2}) \\ Upper\ percentile &= Percentile\ (residual, 100 - \frac{\text{Alpha}}{2\ Gamma}) \end{aligned} \tag{4.9}$$

- Remove the points which their quantities are more than $Upper\ percentile$ and less than $Lower\ percentile$.
- Determine the new value of Alpha.

$$\text{Alpha}_{new} = \text{reduction factor} * \text{Alpha} \tag{4.10}$$

- If $\text{Alpha}_{new} \geq$ Beta, repeat the steps from the computation of the residuals, otherwise halt the loop.

Note that, values of Alpha, Beta, Gamma and Alpha reduction factor are determined by trial and error and can be different from case to case.

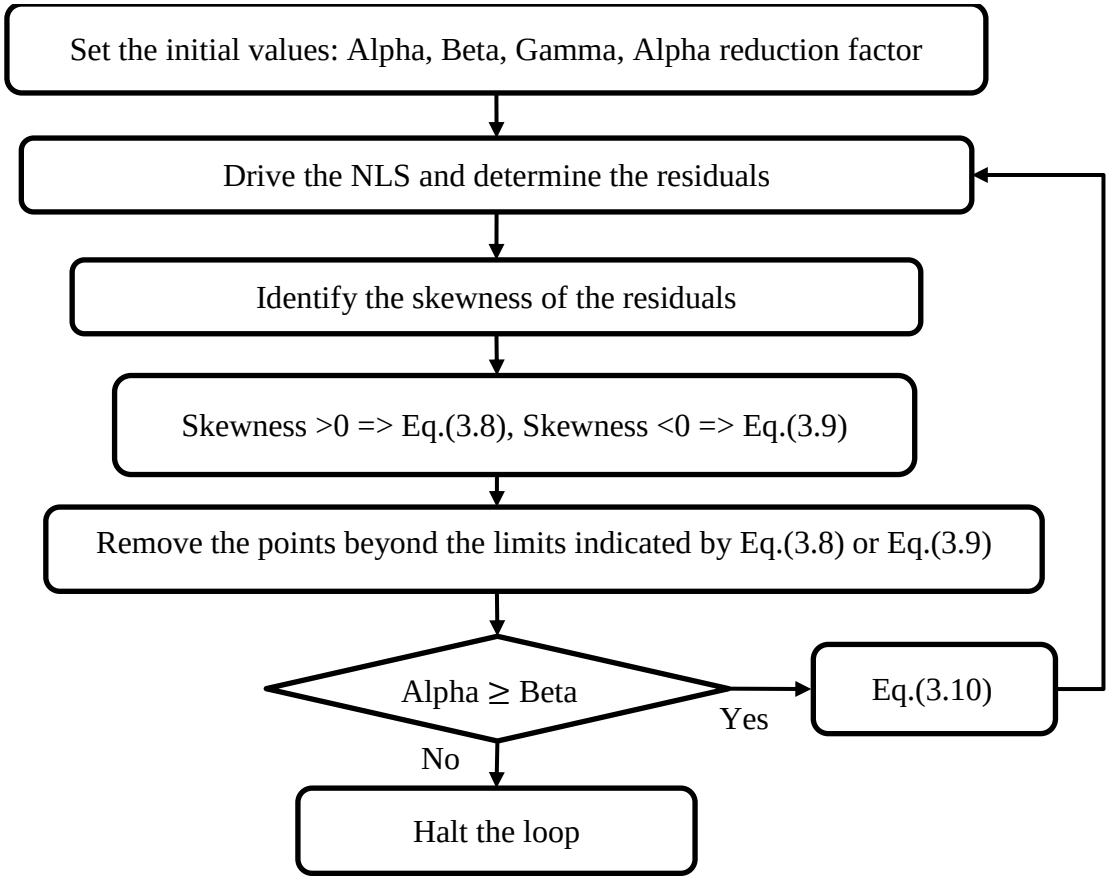

Fig. 4.3. Flowchart of filtering implementation

Fig. 4.4 demonstrates the sequence of implementation of filtering mechanism in the schematic format. The red curvature depicts the target distribution of residual, which presumably there is no information about it. The Gray curvature illustrates the long left tailed distribution of residual. The Blue star and Gray cross represent the percentile of raw data, which is identified as the outlier by the filtering algorithm and the trimmed data, respectively.

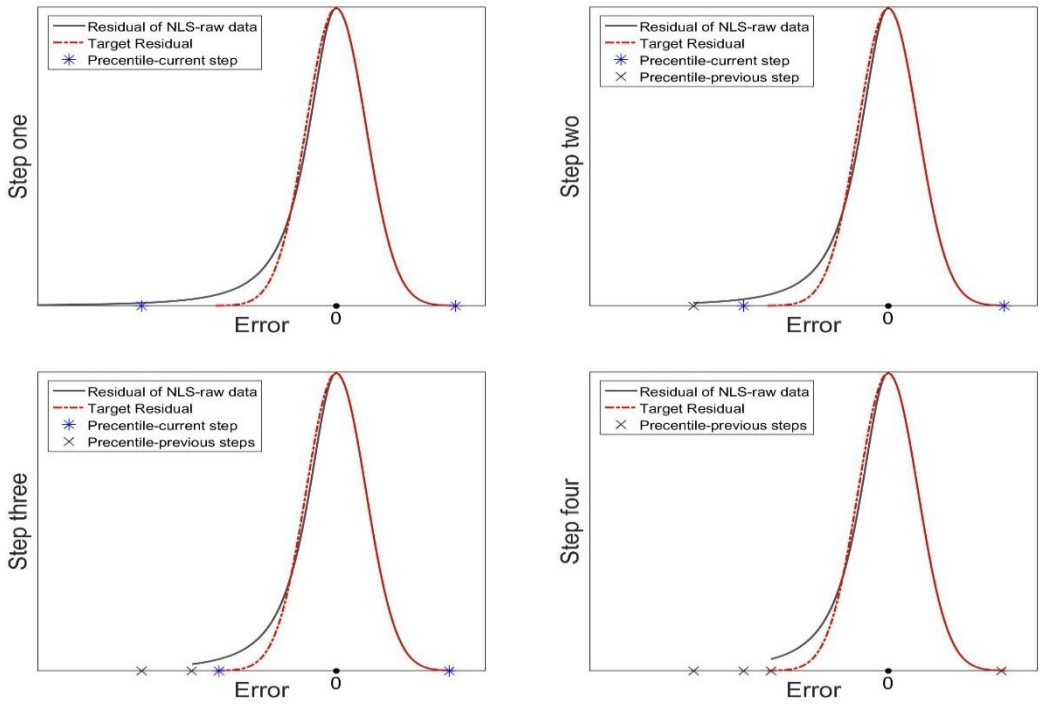

Fig. 4.4. Schematic configuration of residual of NLS and the raw data alongside the target error curvatures according to each skewness reduction step. Respectively from step one, top left to the step four down right

The observed data before (red square) and after (gray star) filtering alongside the NLS optimized curves (dashed and solid lines represent raw and filtered data, respectively) in the normal and log-log scales for Well 1-MHD is demonstrated in Fig. 4.5.

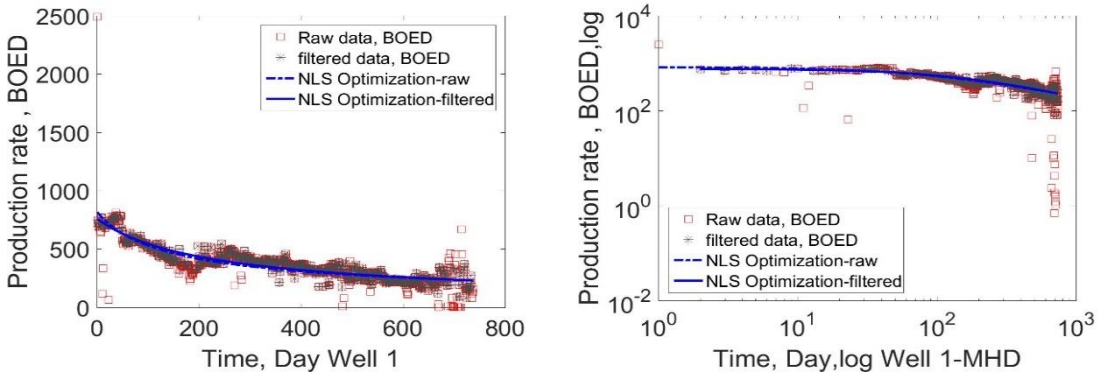

Fig. 4.5. Observed data and nonlinear least square fit before and after filtering

Fig. 4.6 depicts the state of the error between the observed data and optimized curves associated with the raw and filtered data. The boundary lines vividly display the oscillation of the error around zero. The skewness trimming mechanism clearly has put the observed data into more organized order.

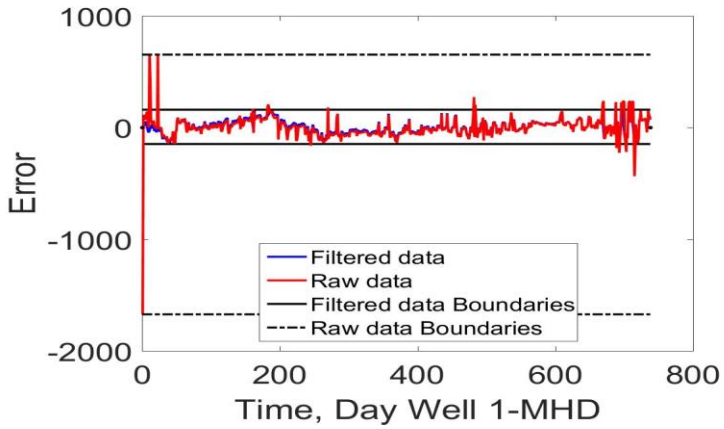

Fig. 4.6. Error between data and optimized curve before and after filtering

A comparison of the raw and observed relative frequency histograms and cumulative density functions are shown in Fig. 4.7. The plot of CDF as well as the relative frequency histogram of raw data illustrates an extreme left and moderate right skewness, which are removed after applying the filtration.

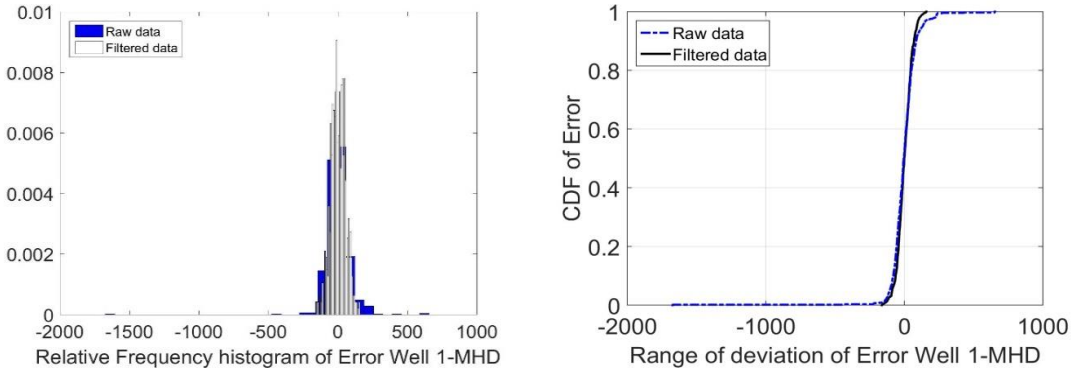

Fig. 4.7. Comparison of the relative frequency histogram and cumulative density function of raw and filtered data

Furthermore, Fig. 4.8 exhibits the Normal distribution fits to the filtered data in the relative frequency histogram and CDF plots. Fig. 4.8 develops the possible shape of the likelihood of Bayesian analysis, which is the trade of between the observed data and forward model's outcome.

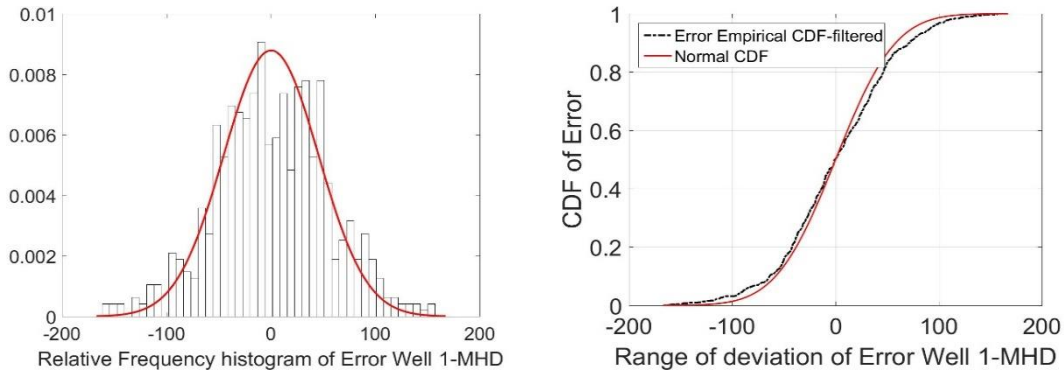

Fig. 4.8. Normal distribution fits to the filtered data, left, relative frequency histogram, right, cumulative density function

4.5. Experimental design and input data

In this research, eight case studies encompass the ensemble of four wells and two forward models are driven. Table 4.2 illustrates the range and initial values of random variables corresponding each forward model. The range is defined to provide a wide constraint for Uniform distribution as the non-informative prior. The initial values and $\sigma_{likelihood}$ are obtained from the NLS optimization of data and statistics of residuals. $\sigma_{likelihood}$ denotes the standard deviation of the likelihood which is set as a constant value for PSAMH.

Applying Eq. (2.18) quantities of the target optimized acceptance rate for MHD and PLED become 0.31 and 0.289, respectively.

Table 4.2 Initial and range of variables in addition to the standard deviation of likelihood of experimental cases

Forward model	Variable	Range of variables [min, max]	MCMC initial values			
			Well 1	Well 2	Well 3	Well 4
MHD	b	[1e-4, 2]	1.38	1e-4	0.726	0.24
	D_i	[1e-4, 20]	0.004	0.0028	0.003	0.0027
	q_i (BOED)	[0.5, 2.5] ×max(data)	757.17	647.85	475.95	514.39
	$\sigma_{likelihood}$	-	45.24	31.93	19.51	34.2
PLED	n	[1e-4, 2]	0.145	0.01	0.11	0.06
	D_i	[1e-4, 20]	0.1	0.928	0.885	0.982
	D_∞	[1e-14, 2e3]	0.0011	0.0027	0.0011	0.0018
	q_i (BOED)	[0.5, 2.5] ×max(data)	1301.3	1698.2	1734.8	1685
	$\sigma_{likelihood}$	-	42.83	31.94	20.39	32.67

Max (data) in Table 4.2 indicates the maximum value of the observed data in each case.

4.6. MCMC implementation

We implemented MCMC experiment, firstly by selecting the initial values of random variables, recognition of the plausible shape of the likelihood and determination of priors. To achieve the initial values and associated distribution of likelihood along with the quantity of standard deviation, the NLS optimization is employed (Fig. 4.8). The number of concurrent chains for both methods is initially set to $m = 100$ and after 1e4 iteration is

reduced to $m' = 30$. The values of iteration for MHD and PLED are set to 1e6 and 2e6, respectively.

In MCMC, it is a common practice to set a point (burn-in), which indicates the state of stationary. The cumulative mean and standard deviation (sometime they are referred as the online mean and standard deviation) besides the general trend of MCMC experiment for each variable are the methods which are applied to identify the burn-in point. The burn-in point eventually is set to 20% of total iterations for either forward models.

The Bayesian calibration implementation, subsequently, is merely delineated through the comprehensive description of cases Well 1-MHD and Well 1-PLED and the associated results of entire cases are provided in the section 0.

4.6.1.1. Well 1-MHD

MCMC experiments

MCMC experiments of three parameters of MHD model, which are set as random variables are demonstrated in Fig. 4.9, left. The Cumulative mean and standard deviation, which are initiated by C.M. and C.STD, demonstrate the state of stationary of MCMC experiment, Fig. 4.9, right.

MCMC experiment as well as the cumulative mean and standard deviation reveal the rapid convergence of MCMC chain. The stationary is presumably achieved when both cumulative mean and standard deviation become straight lines.

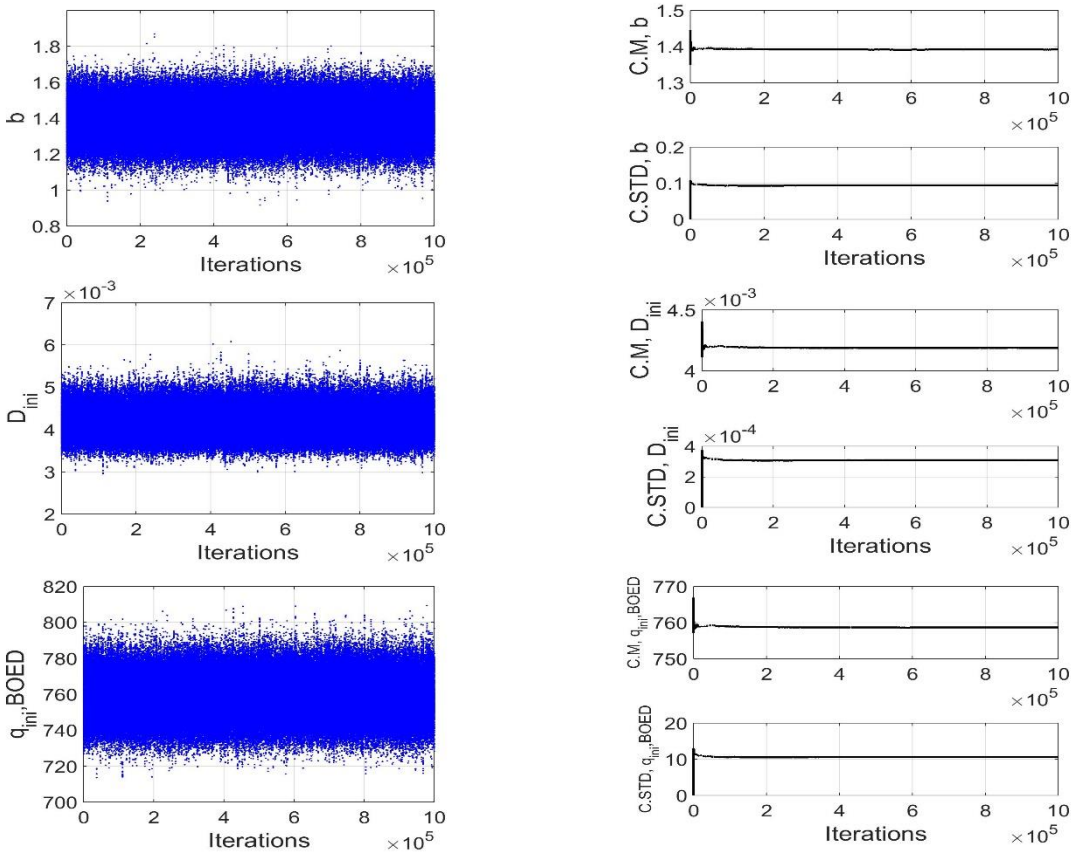

Fig. 4.9. MCMC experiments and the corresponding cumulative mean and standard deviation of MHD

Relative frequency histogram and cumulative density function

The vector of random samples after burn-in point construct the posterior space. Obtaining the posterior space, one would be able to infer the statistics. The relative univariate frequency histogram (R.F.H), joint distribution and CDF plots of random variables are demonstrated in Fig. 4.10.

The formation of joint distribution plots unveils the strong positive correlation between random variables, whilst the red zones at the center represent the higher probability of occurrence in compare to the edges. The relative frequency histograms and CDF plots also present some degree of tendency toward the Normal distribution.

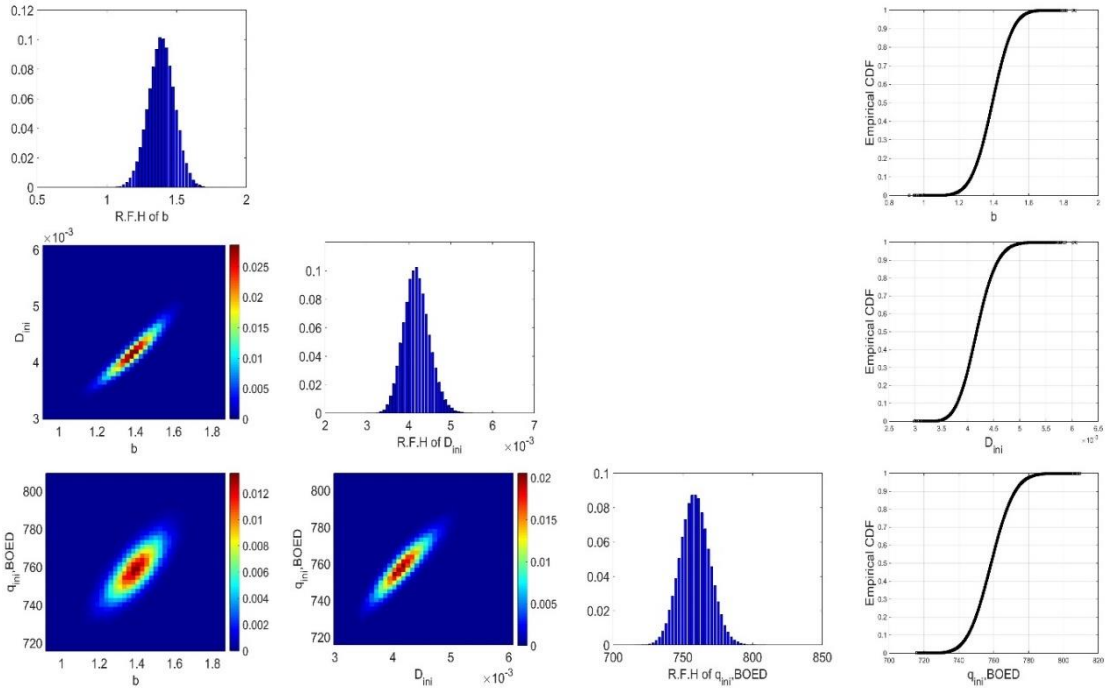

Fig. 4.10. Joint distribution, relative frequency histogram and CDF of MHD model random variables

Three-dimensional joint distribution

Fig. 4.11 also displays the joint distribution of MHD random variables in an oval configuration. The rad region at the center of oval volume indicates the higher probability of occurrence.

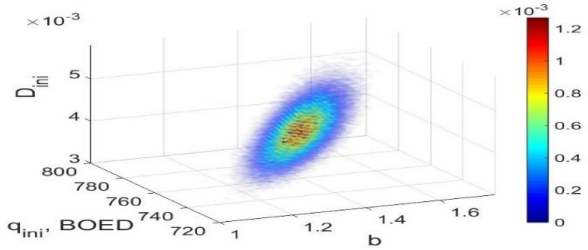

Fig. 4.11. Three-dimensional joint distribution of MHD random variables

Realizations

It is evident that, one of the main advantages of the Bayesian analysis is the ability of prediction of model behavior by drawing thousands of random samples from the posterior

and plugging them into the forward model. The whole procedure appears as the realization in the Bayesian literature. In the subsequent section, we deliver Bayesian inference through the 10,000 realizations of forward models.

Current production time (daily basis)

Likely realizations corresponding the current production time in compare to the observed data (red stars) in normal and log-log scales are displayed in Fig. 4.12, top right and left. The mean of realizations, furthermore, demonstrate a perfect match to the observed data (Fig. 4.12, middle, left and right). The standard deviation of model realizations, which is occasionally interpreted as the level of confidence or inference certainty, are depicted in Fig. 4.12, down, left and right.

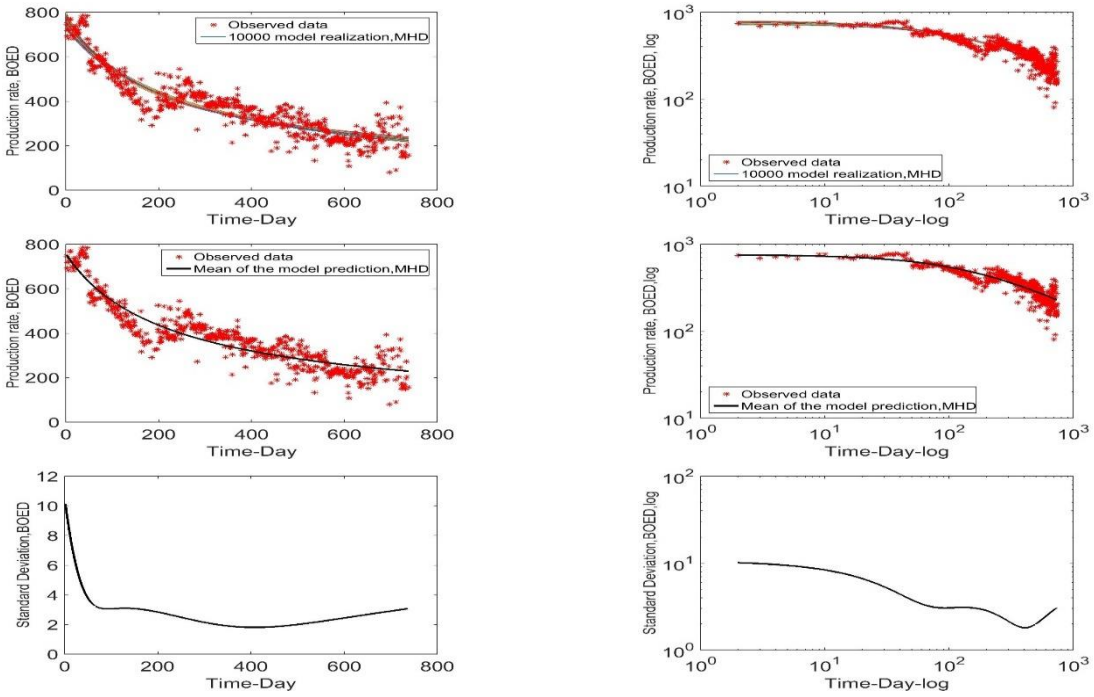

Fig. 4.12. 10,000 realizations, mean and standard deviation of realization in normal and log-log scales

The plot of standard deviation connotes that at the beginning of the process, the level of confidence is lower than the other times. By proceeding, the certainty gradually sours whilst at the end it is slightly altered.

Current production time (cumulative)

Another significance of Bayesian model realization emerges from the computation of the cumulative production. In this research, we applied the principles of the numerical integration to achieve the cumulative production out of the forward model functions.

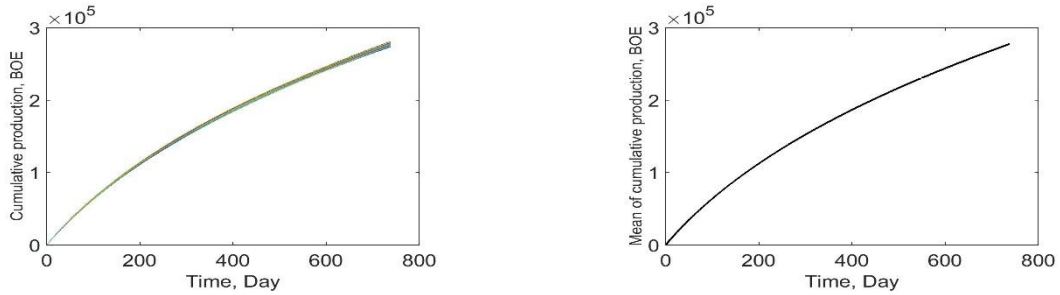

Fig. 4.13. Left, cumulative production and right, mean of, model realizations

The dispersion of model realizations is soared by proceeding in time (Fig. 4.13, right)

30 years' prediction of production time (daily basis)

In order to provide some insights about the future production's status of a well, by supplying the model realization method, the current production data extrapolates to 20 or 30 years of productions. Note that, empirical models sometimes underestimate or overestimate the long-term production; hence taking precautions are advisable concerning the type of forward models. Fig. 4.14 demonstrates the 30 years production realization as well as the mean and standard deviation of realizations in both normal and log-log scaled.

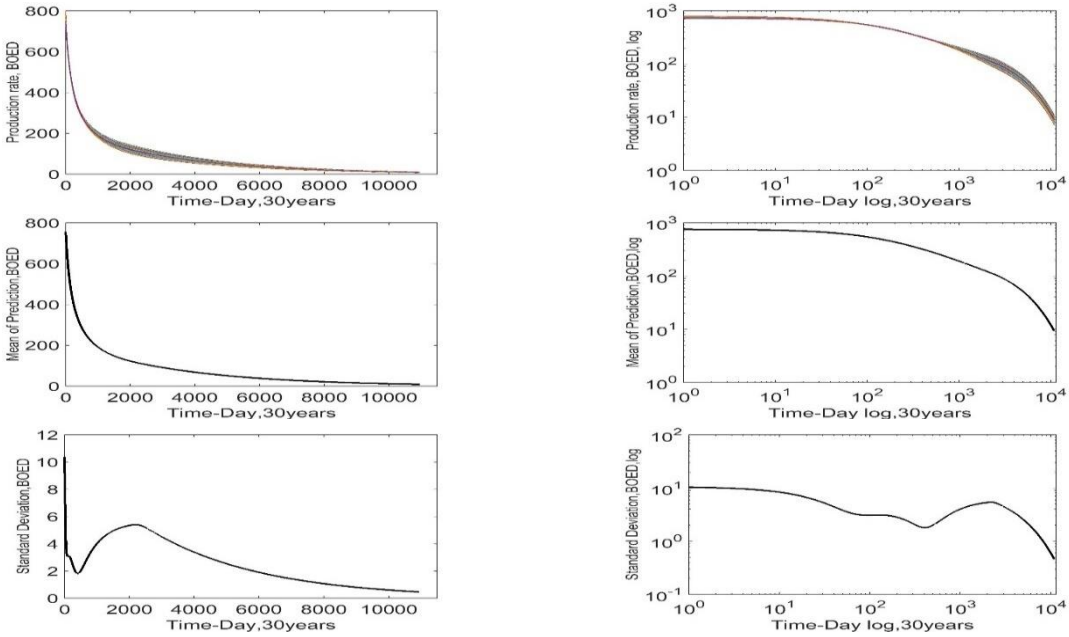

Fig. 4.14. 30 years realization, mean and standard deviation of model realization in both normal and log-log scales

Considering the plot of standard deviation of realizations, it is alluded that by increasing the production time up to 2000 days, the level of certainty is declined and after that, it is continuously improved.

30 years' prediction of production time (cumulative)

Similar to the current time production, the 30 years' cumulative realization is also provided (Fig. 4.15). However, the diffusion of realization is considerably higher than the current time production.

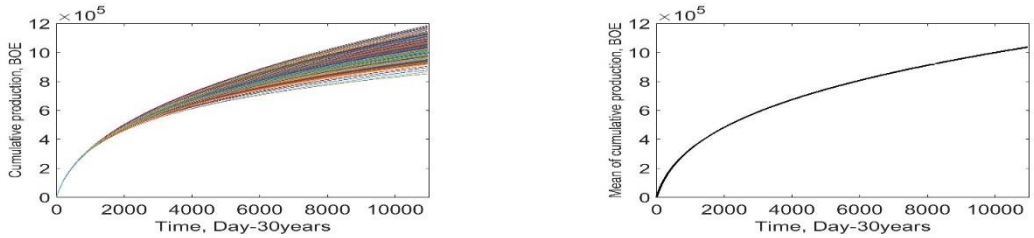

Fig. 4.15. Left, 30 years and right, mean of cumulative realizations

4.6.1.2. Well 1-PLED

The entire aforementioned procedures about MHD are also held for PLED model. Nevertheless, PLED comprises one more random variable in compare to MHD and hence is necessitated to run to more iterations (2e6). Thus, MCMC computationally takes more time.

MCMC experiments

MCMC experiments and cumulative means and standard deviations of PLED random variables are shown in Fig. 4.16. Despite the immediate convergence of MCMC, the pattern of sampling is clearly changed in comparison to the MHD.

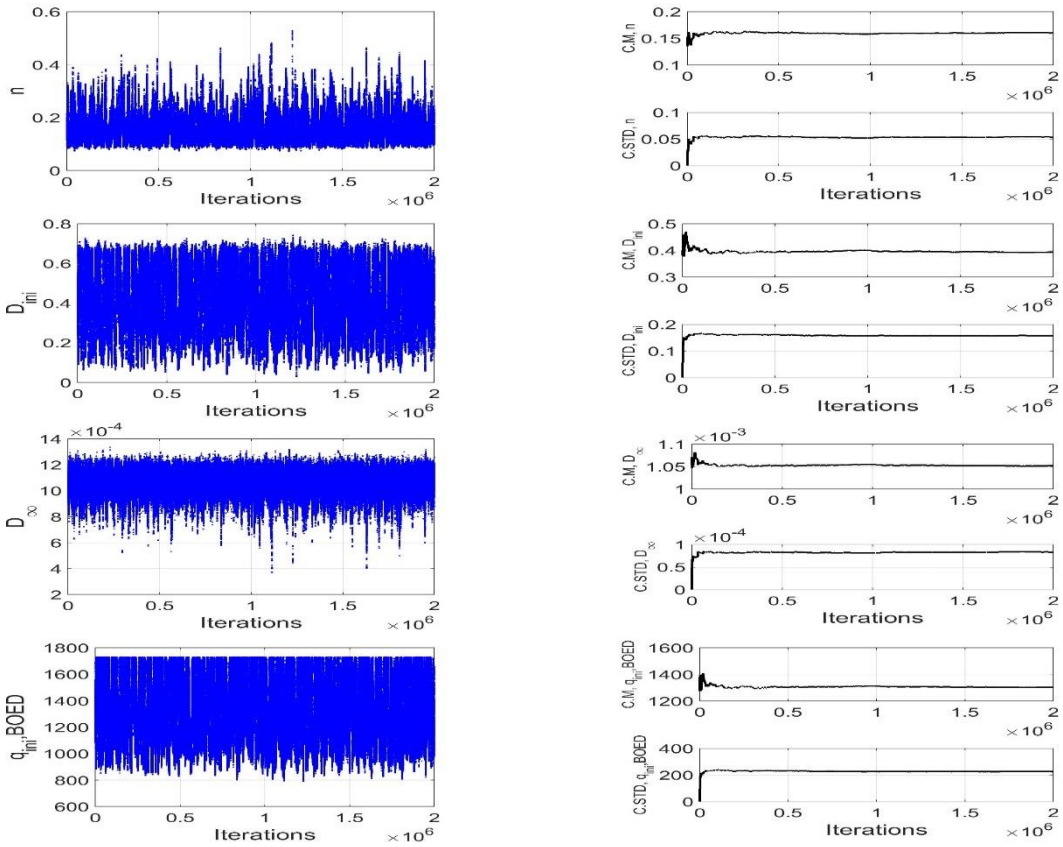

Fig. 4.16. MCMC experiments and the associated cumulative mean and standard deviation of PLED

Relative frequency histogram and cumulative density function

The univariate relative frequency histograms along with the joint distribution of random variables for PLED, distinctly imply the irregularity in the structure of the posterior in compare to the MHD (Fig. 4.17).

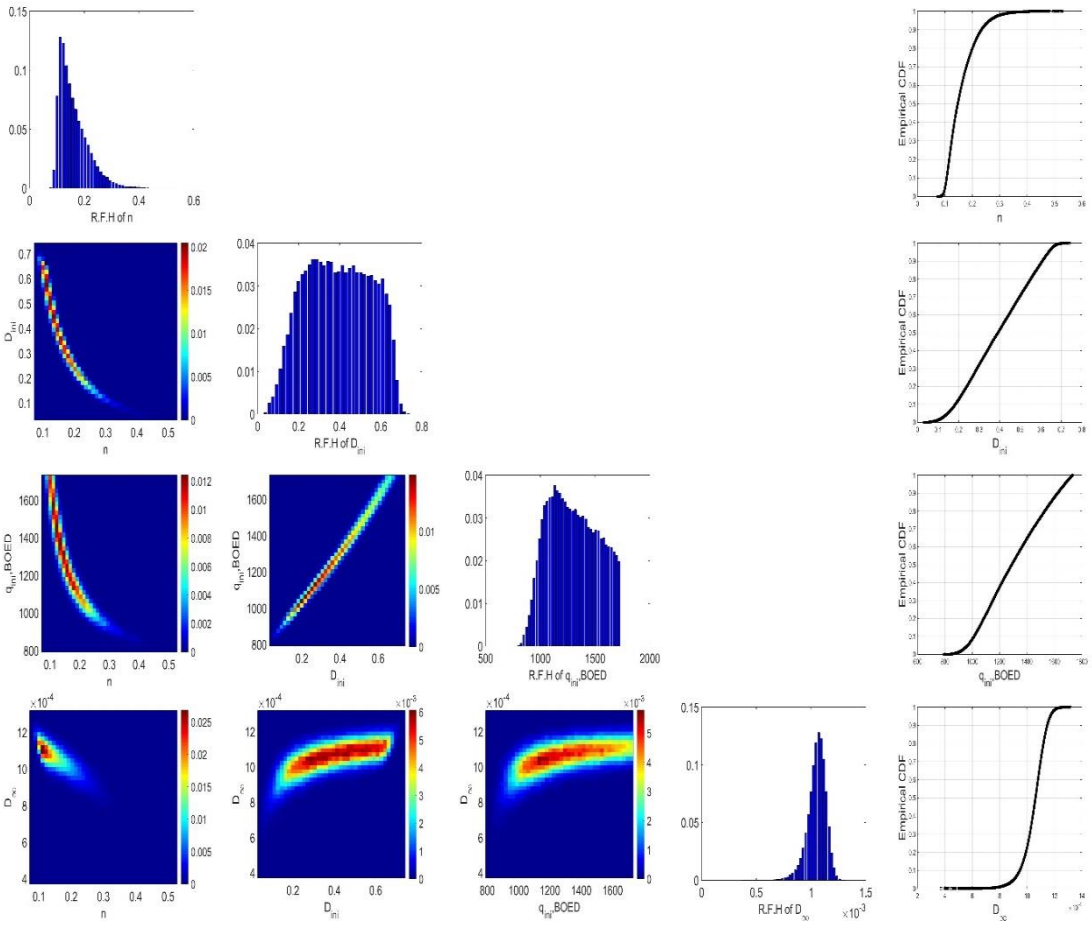

Fig. 4.17. Joint distribution, relative frequency histogram and CDF of PLED random variables

Generally, on contrary to the MHD, the correlation coefficient of variables in PLED appears in both negative and positive quantities.

Realizations

Current production time (daily basis)

Realization of PLED in both normal and log-log scales are matched to the observed data, Fig. 4.18. In addition, the mean and standard deviation of model realization are retrieved and illustrated in the same figure.

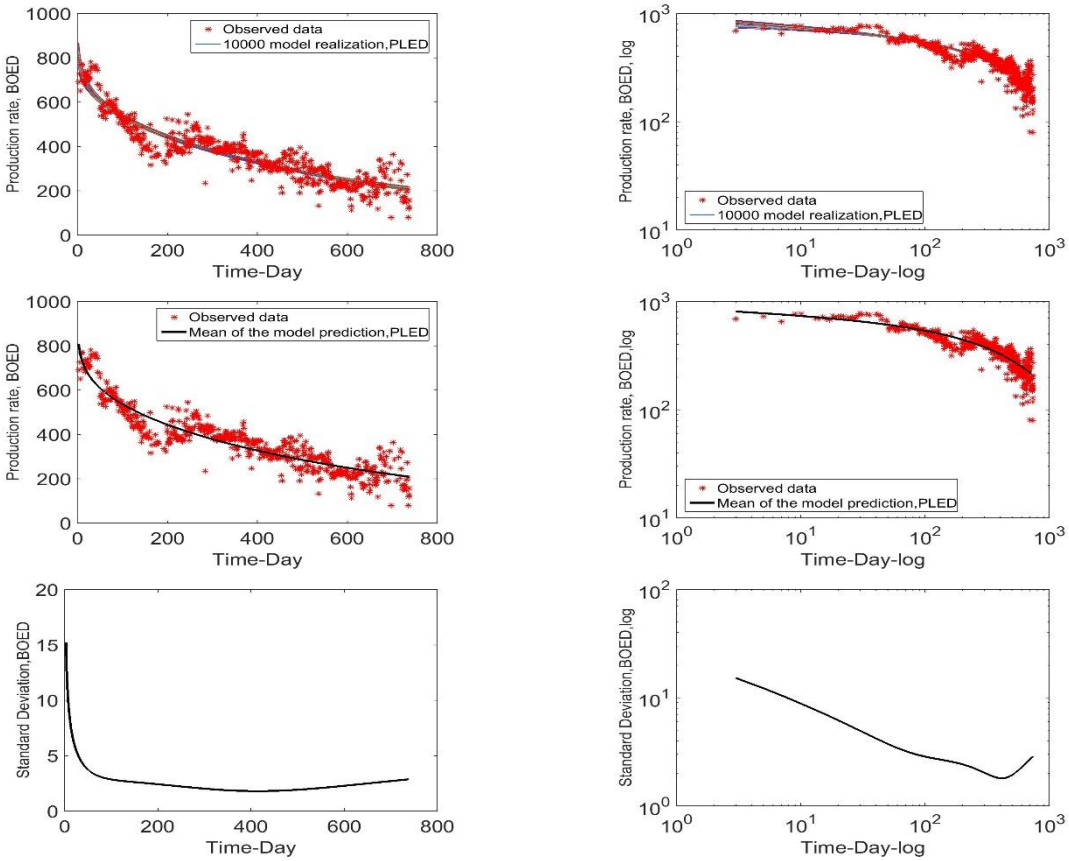

Fig. 4.18. 10,000 realizations, mean and standard deviation of realization in normal and log-log scales, PLED model

The mean of realization presents a perfect match to the observed data. Moreover, as it is expected, the standard deviation begins with the larger quantity and accepts smaller values by increasing the production time.

Current production time (cumulative)

Realizations of the cumulative production of current time elucidates more diffusion at the end of current production time (Fig. 4.19).

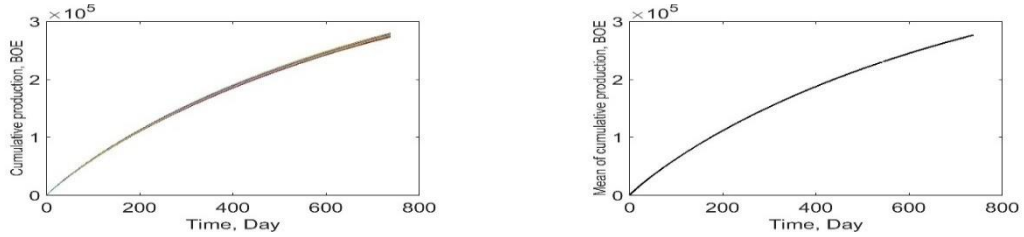

Fig. 4.19. Left, cumulative production and right, mean of, model realizations

30 years' prediction of production time (daily basis)

30 years realization of model prediction, the mean and standard deviation of realizations plots in PLED depict similar behavior as MHD (Fig. 4.20). However, the level of confidence is slightly less than MHD model at the beginning.

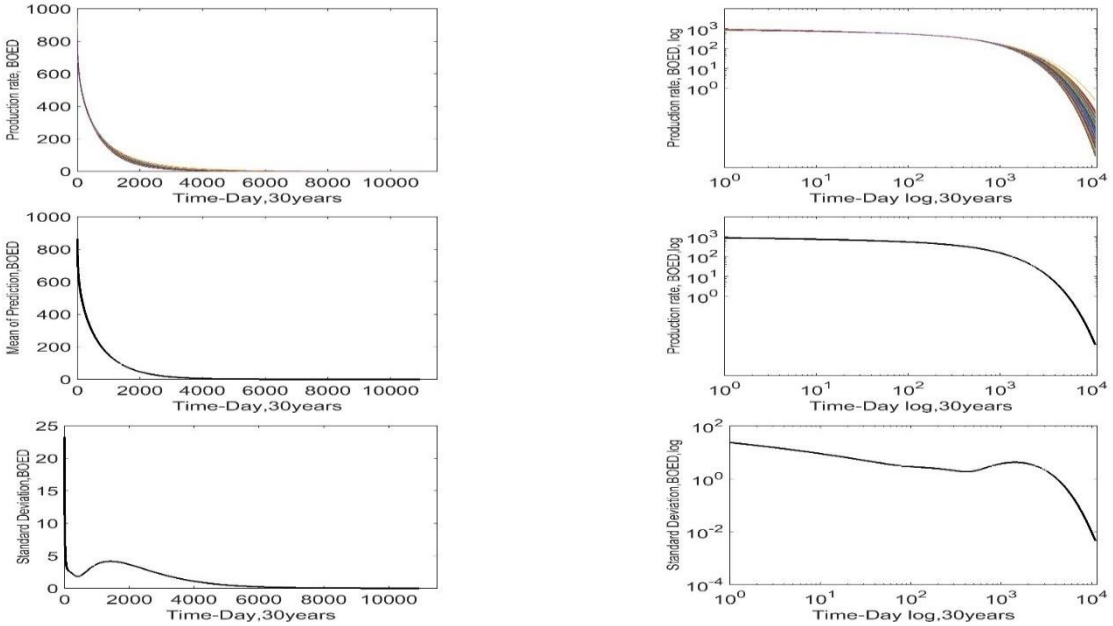

Fig. 4.20. 30 years realization, mean and standard deviation of model realization in both normal and log-log scales

30 years' prediction of production time (cumulative)

Fig. 4.21 demonstrates the cumulative production for 30 years. The trend of production after 2000 days almost appears to be constant. Some deviation can be seen in the developed configuration, although it is far less than MHD.

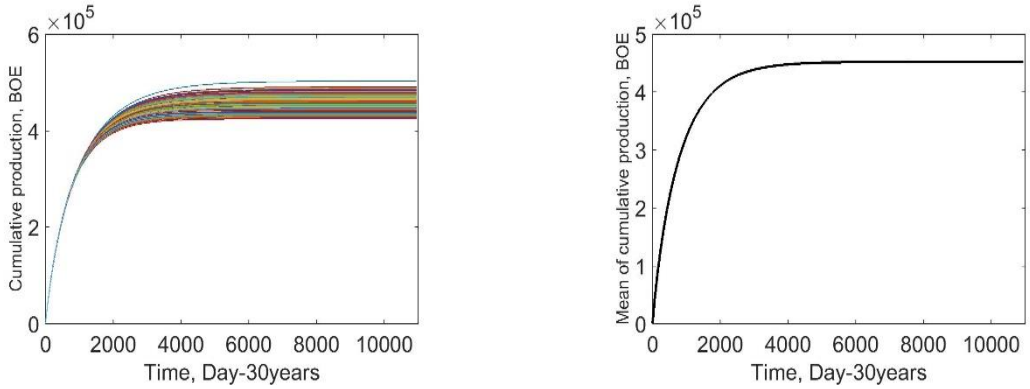

Fig. 4.21. Left, 30 years and right, the mean of cumulative realizations

4.7. Results and discussion

Following, the retrieved values of mean and standard deviation of forward models' random variables as well as the acceptance rate of experimental cases are provided in Table 4.3. The captured acceptance rates produced a great degree of agreement to the optimized acceptance rates. Nevertheless, the general trend of variables is altered from one well to another. For instants, variable b of MHD extended from 0.0068 in Well 2 to 1.39 in Well 1, while n covered a range between 0.0039 and 0.16. Parameter D_∞ provided small quantities in all cases. Variable q_i presented no consistency in MHD and PLED. q_i is accepted noticeable larger mean and standard deviation values in PLED.

Table 4.3 Mean and standard deviation of model varaiables along with the acceptance rate of MCMC experiments

Forward model	Variable	Well 1		Well 2		Well 3		Well 4	
		μ	σ	μ	σ	μ	σ	μ	σ
MHD	b	1.4E0	9.4E-2	6.8E-3	9.4E-3	7.3E-1	4.9E-2	0.25	5.4 E-2
	D_i	4.2E-3	3.1E-4	2.8E-3	3.0E-6	3.1E-3	1.0E-4	2.7E-3	1.0 E-4
	q_i (BOED)	758.7	10.6	648.5	3.56	476.3	4.1	514.7	4.9 9
	Acceptance rate %	31.03		31.02		31.14		30.99	
PLED	n	1.6E-1	5.3E-2	3.9E-2	1.2E-1	1.3E-1	1.6E-2	7.5E-2	1.6 E-2
	D_i	3.9E-1	1.7E-1	2.6E-1	3.1E-1	7.4E-1	1.1E-1	7.8E-1	1.6 E-1
	D_∞	1E-3	8.0E-5	2.7E-3	2.3E-4	1.1E-3	5E-5	1.8E-3	5E-5
	q_i (BOED)	1301.9	223.6	896.9	314.1	1505.3	161.3	1386.5	206 .6
	Acceptance rate %	29.1		28.73		28.99		28.95	

4.7.1. Comparison of MHD and PLED

Comparison plots of all experimental cases are demonstrated and discussed here, to be better able to perceive the behavior of forward models regarding the different production data.

4.7.1.1. Well 1

Current production time (daily basis)

The mean of realization in both models show a good match with the observed data. However, they crossed each other in several times (Fig. 4.22). The plot of standard deviation also exhibits a good agreement between both methods.

Current production time (cumulative)

Both methods present the same cumulative production mean and standard deviation for the current time (Fig. 4.23).

A slight divergence between the cumulative standard deviation of models introduced more confidence in MHD.

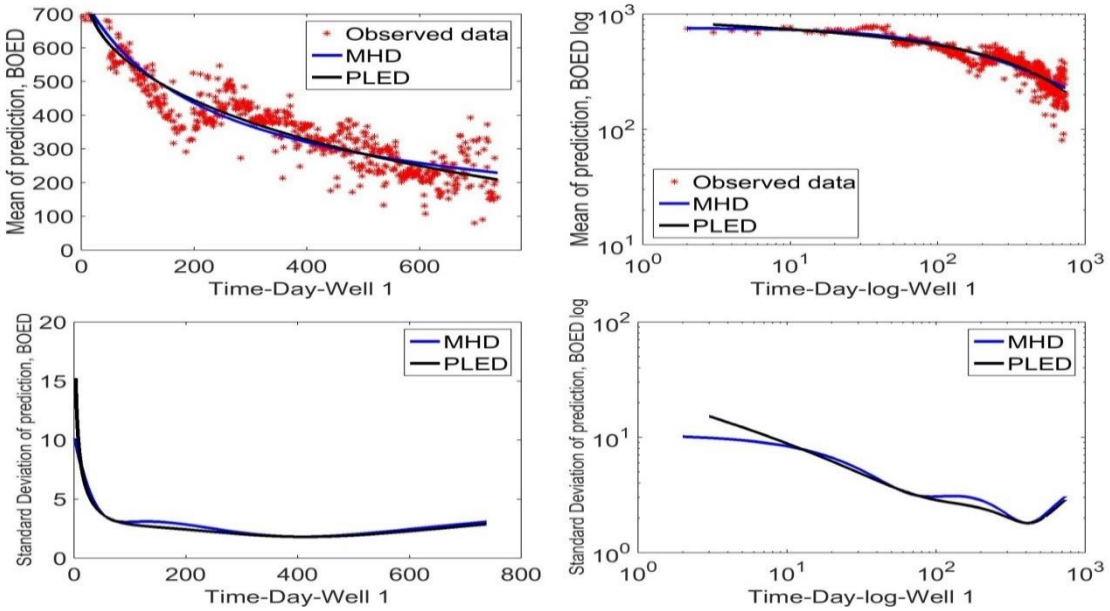

Fig. 4.22. Mean and standard deviation of MHD and PLED in normal and log-log scaled, Well 1

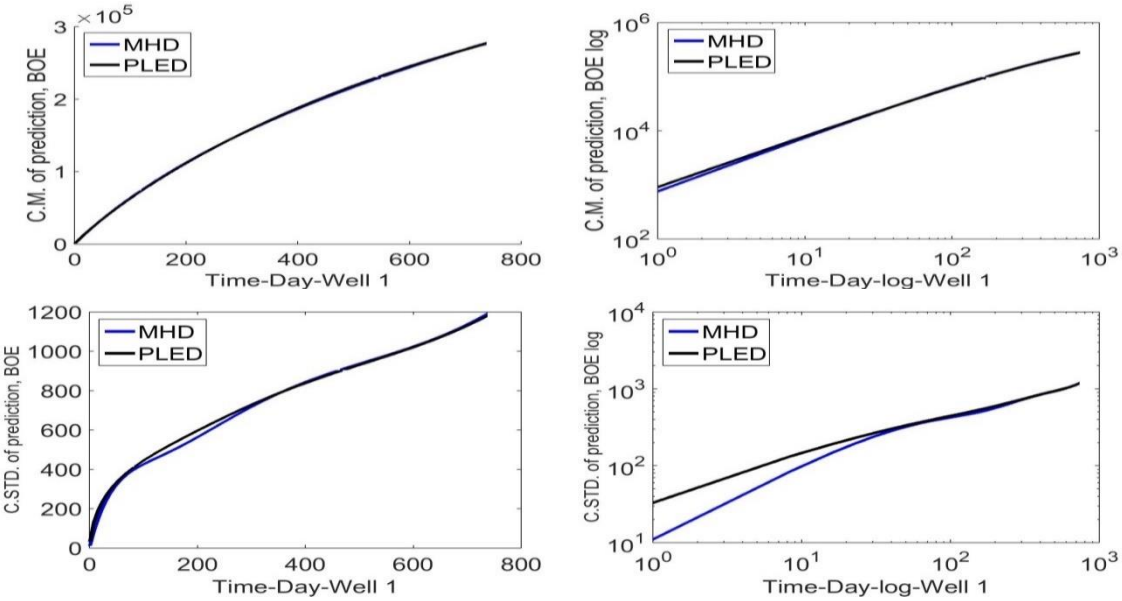

Fig. 4.23. Mean and standard deviation of the cumulative production of MHD and PLED in normal and log-log scales, Well 1

30 years' production time (daily basis)

PLED in Fig. 4.24 depicts less production for a long period of time, although, the corresponding standard deviation becomes less which means more confidence about the modeling by proceeding in time.

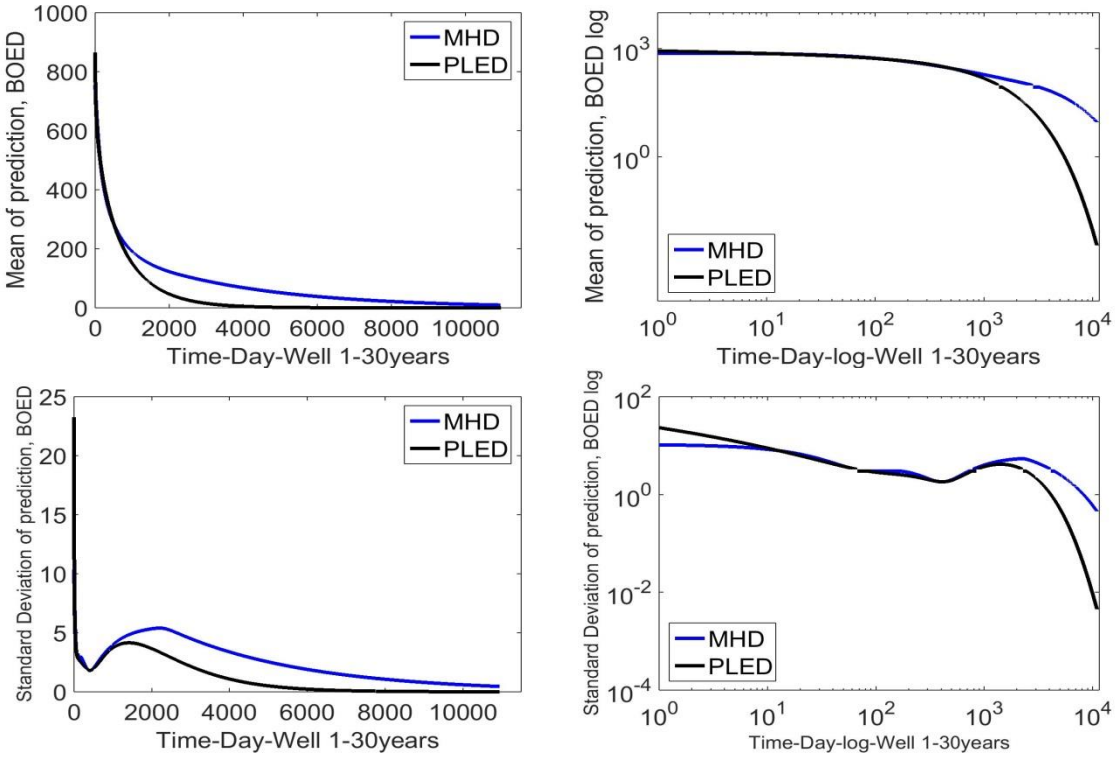

Fig. 4.24. Mean and standard deviation of 30 years production for MHD and PLED in normal and log-log scales, Well 1

30 years' production time (Cumulative)

The cumulative production also pursues the same trend. Although, the cumulative mean of MHD provides more production during 30 years, the level of certainty is relatively less than PLED (Fig. 4.25).

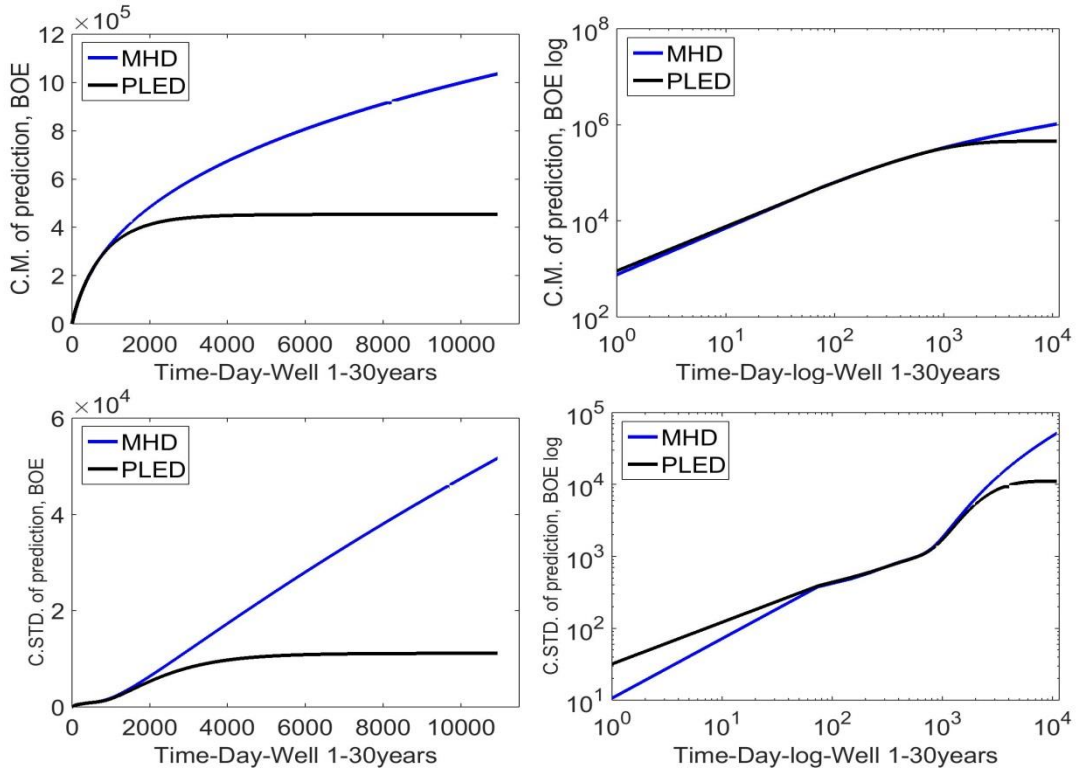

Fig. 4.25. Mean and standard deviation of 30 years cumulative production for MHD and PLED in normal and log-log scales, Well 1

4.7.1.2. Well 2

Current production time (daily basis)

The mean of realizations in both methods entirely cover each other (Fig. 4.26). The standard deviation of MHD, however, started with smaller quantity but after 300 days, both methods coincide each other.

Current production time (cumulative)

The difference between two methods is not distinguishable in Fig. 4.27. Nevertheless, a trivial departure in the standard deviation is captured.

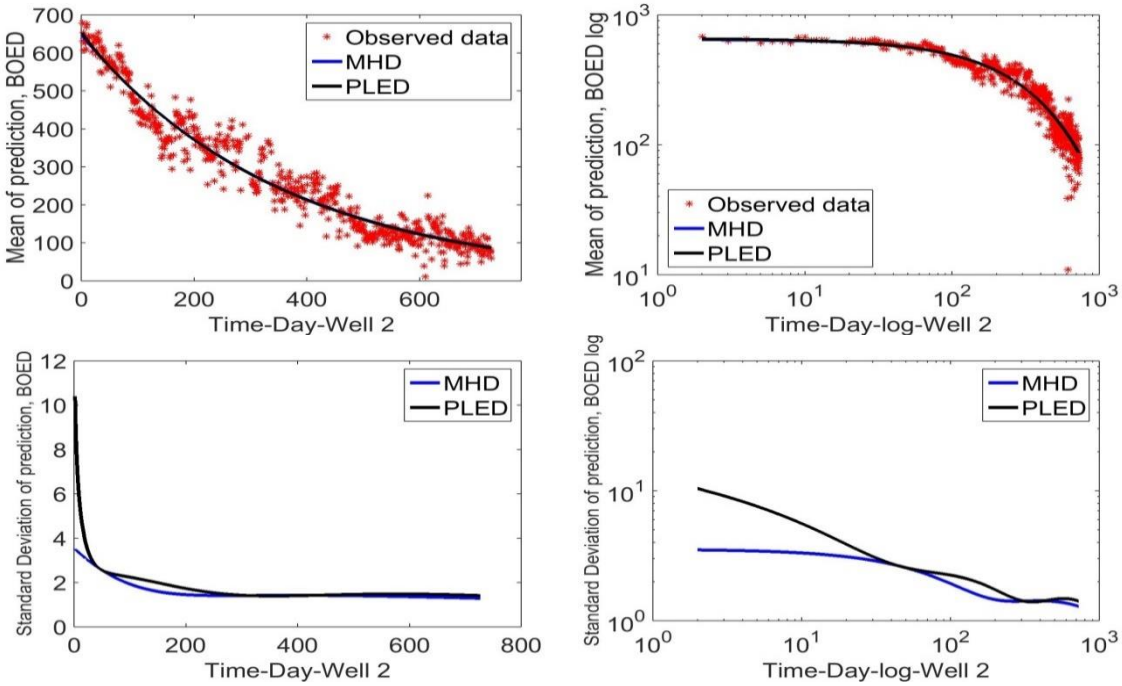

Fig. 4.26. Mean and standard deviation of MHD and PLED in normal and log-log scales, Well 2

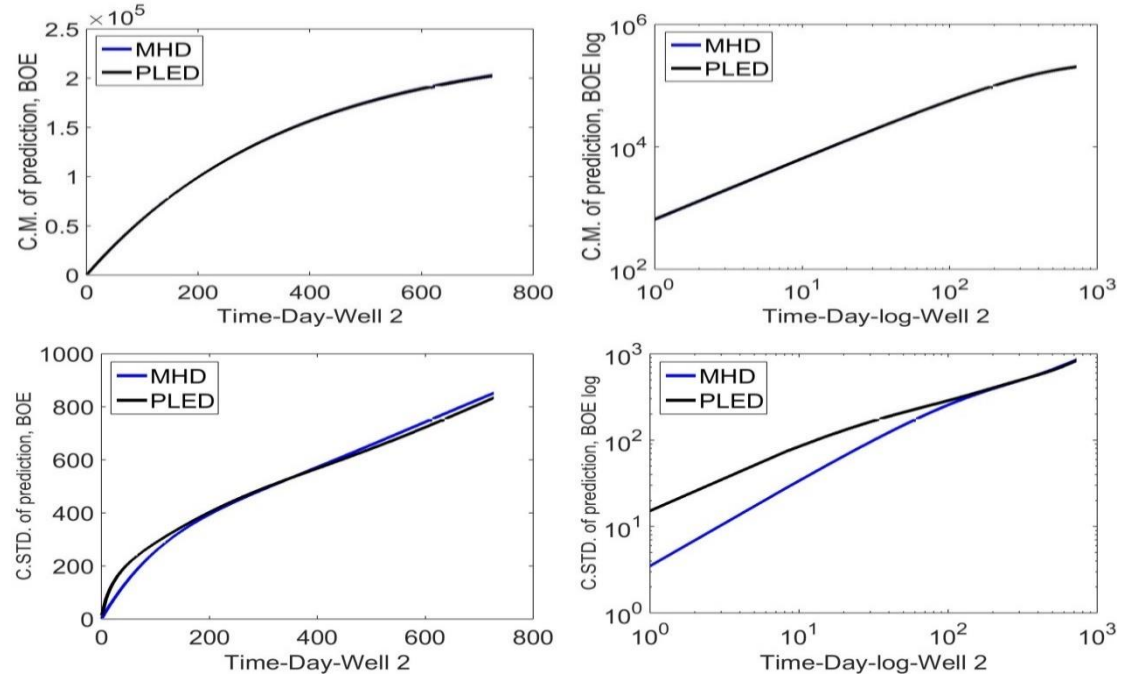

Fig. 4.27. Mean and standard deviation of the cumulative production of MHD and PLED in normal and log-log scales, Well 2

30 years' production time (daily basis)

The agreement in the production is continued into 30 years of prediction (Fig. 4.28). In the normal plots, no significance can be recognized while the log-log scale provides a better assessment regarding the productions.

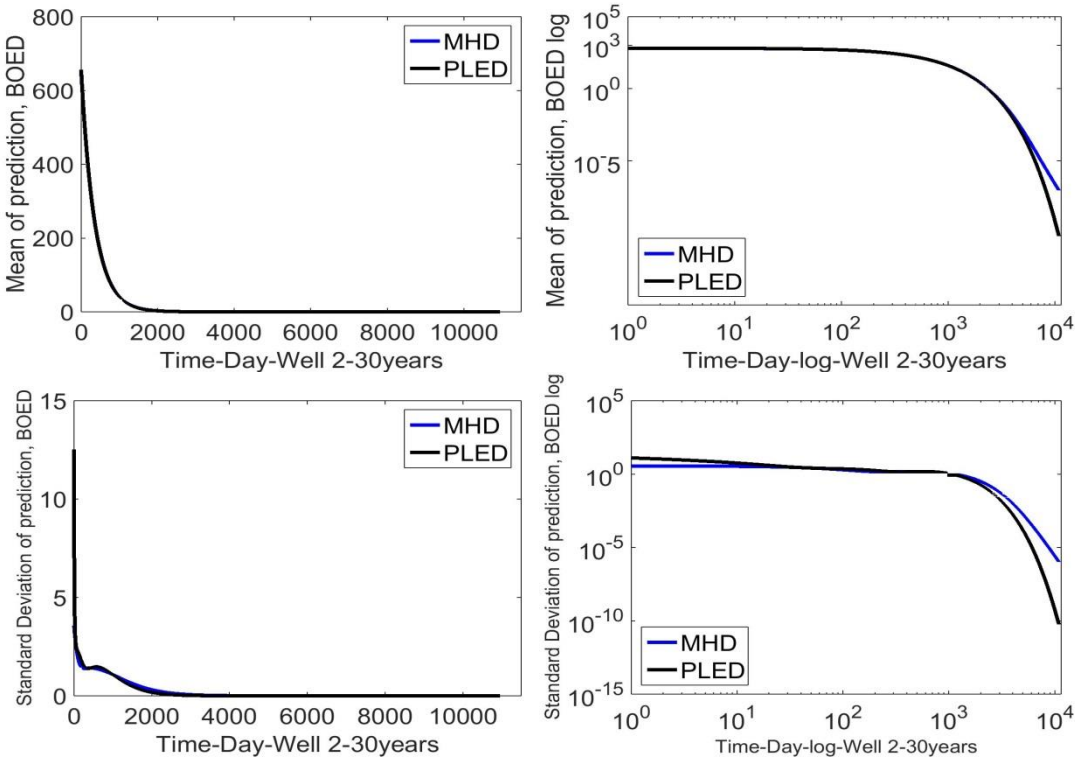

Fig. 4.28. Mean and standard deviation of 30 years production for MHD and PLED in normal and log-log scales, Well 2

30 years' production time (Cumulative)

The only noticeable difference in the cumulative production for 30 years appears in Fig. 4.29 regarding standard deviations when the level of confidence become less in the MHD model.

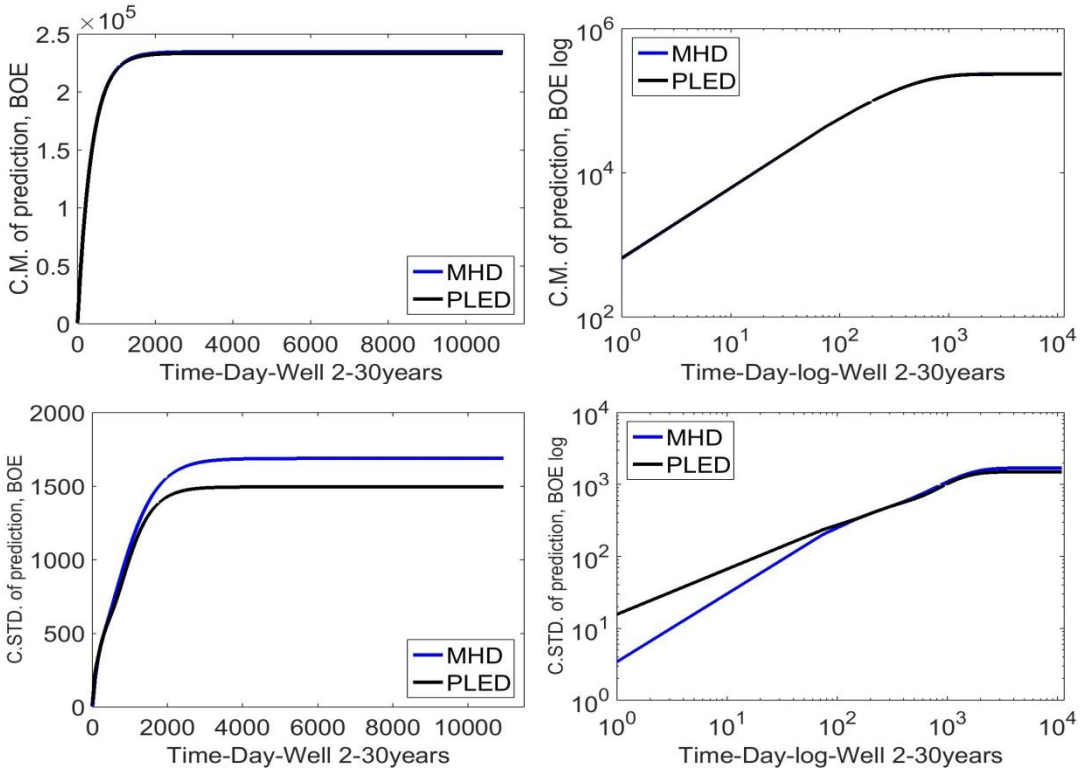

Fig. 4.29. Mean and standard deviation of 30 years cumulative production for MHD and PLED in normal and log-log scales, Well 2

4.7.1.3. Well 3

Current production time (daily basis)

The mean and standard deviation of realization of both methods coincide with each other and a slight difference can be observed for the current production time.

Current production time (cumulative)

The mean of cumulative production depicts similar trend; however, some deviation can be recognized in the plot of standard deviation. The certainty of PLED method on contrary to the earlier cases moderately becomes less than MHD.

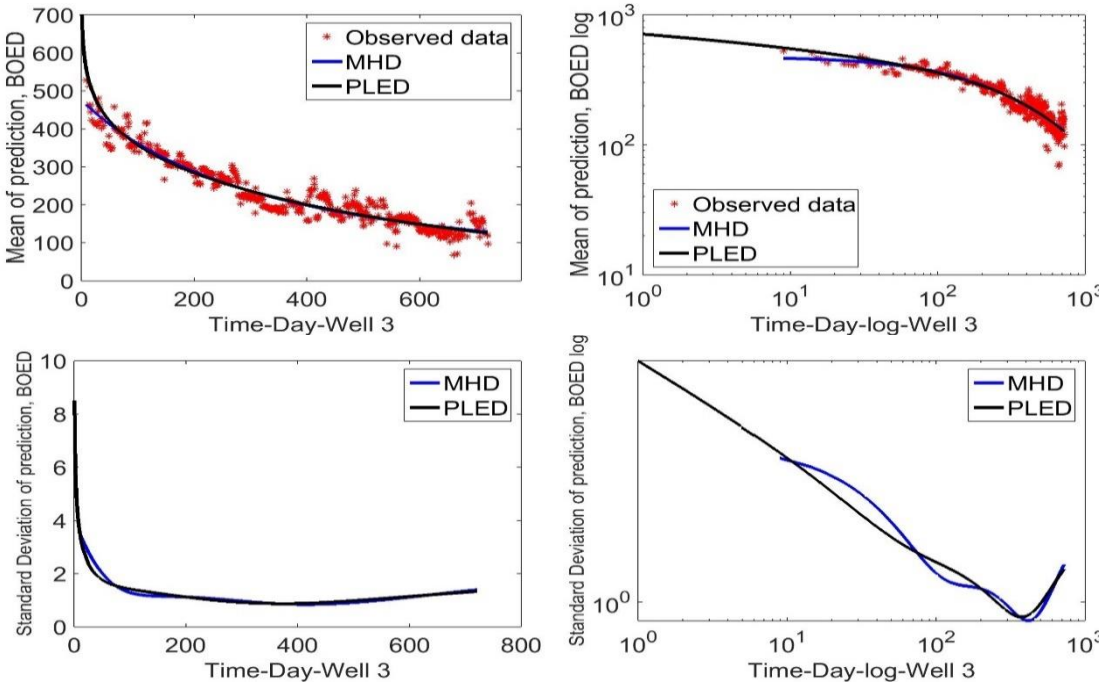

Fig. 4.30. Mean and standard deviation of MHD and PLED in normal and log-log scales, Well 3

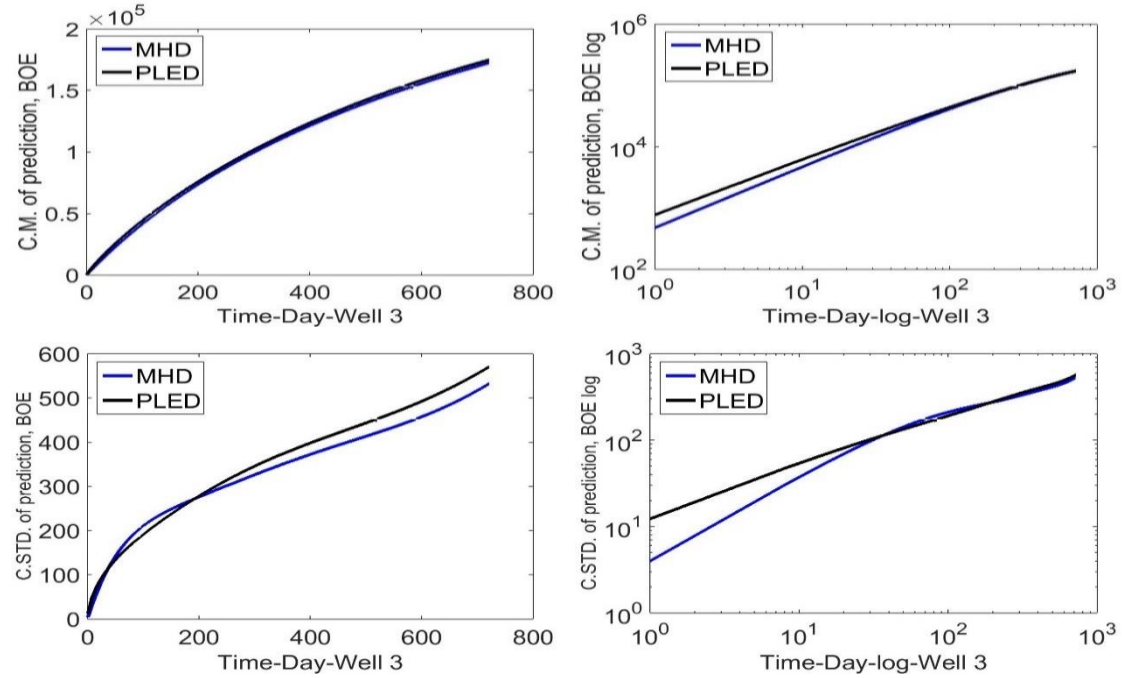

Fig. 4.31. Mean and standard deviation of the cumulative production of MHD and PLED in normal and log-log scales, Well 3

30 years' production time (daily basis)

The comparison plots of mean of 30 years' production depicts an extra production in MHD over PLED. Meanwhile, the level of confidence in MHD gradually declines after a certain production time.

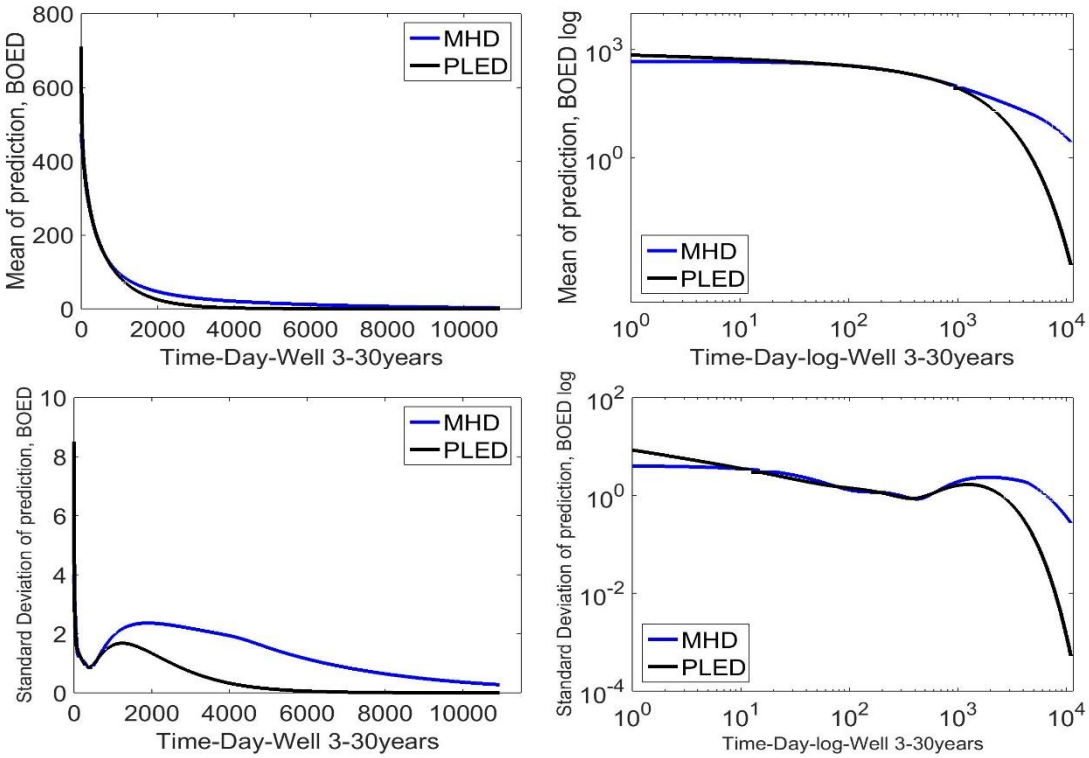

Fig. 4.32. Mean and standard deviation of 30 years production for MHD and PLED in normal and log-log scales, Well 3

30 years' production time (Cumulative)

Fig. 4.33 displays the diffusion between the cumulative mean and standard deviation of realization during 30 years of production. MHD provides a developed production while PLED shows less progress after about 2000 days. In addition, MHD demonstrates less certainty regarding the production time.

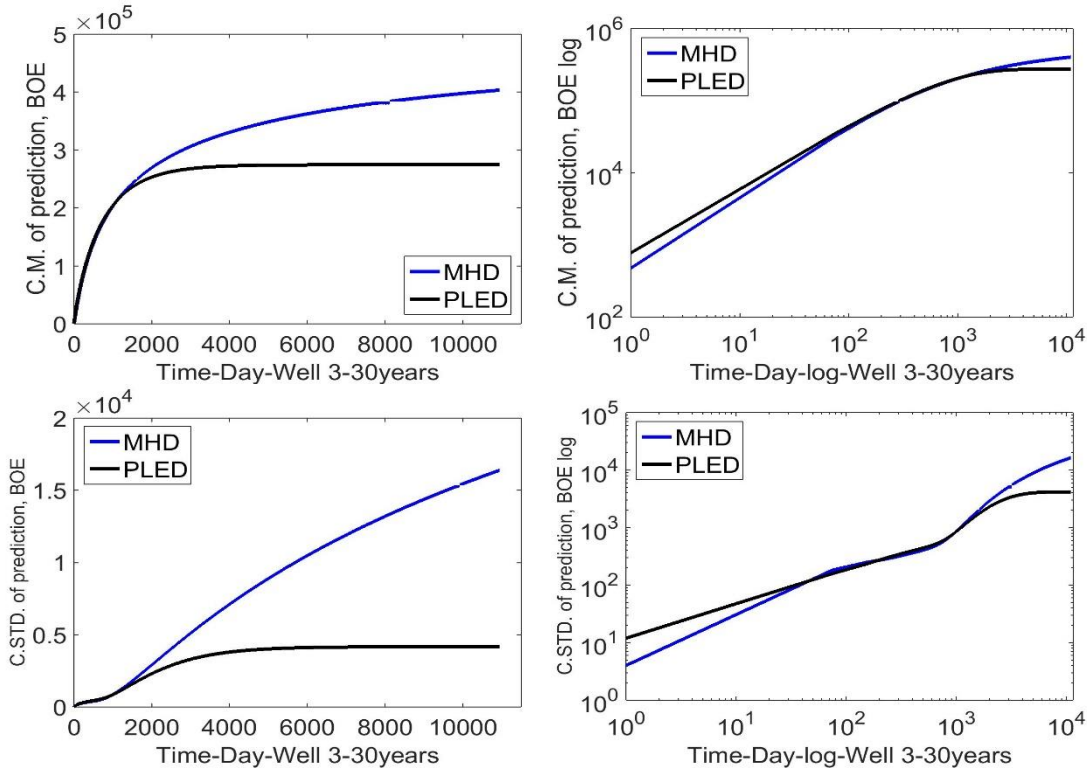

Fig. 4.33. Mean and standard deviation of 30 years cumulative production for MHD and PLED in normal and log-log scales, Well 3

4.7.1.4. Well 4

Current production time (daily basis)

The mean of realizations of both methods features an acceptable match with observed data. The plot of standard deviation almost present similar trend for MHD and PLED.

Current production time (cumulative)

The plot of cumulative production in MHD and PLED pursues the same behavior for daily production. In this case either methods provide similar results associated with the mean and standard deviation of realizations.

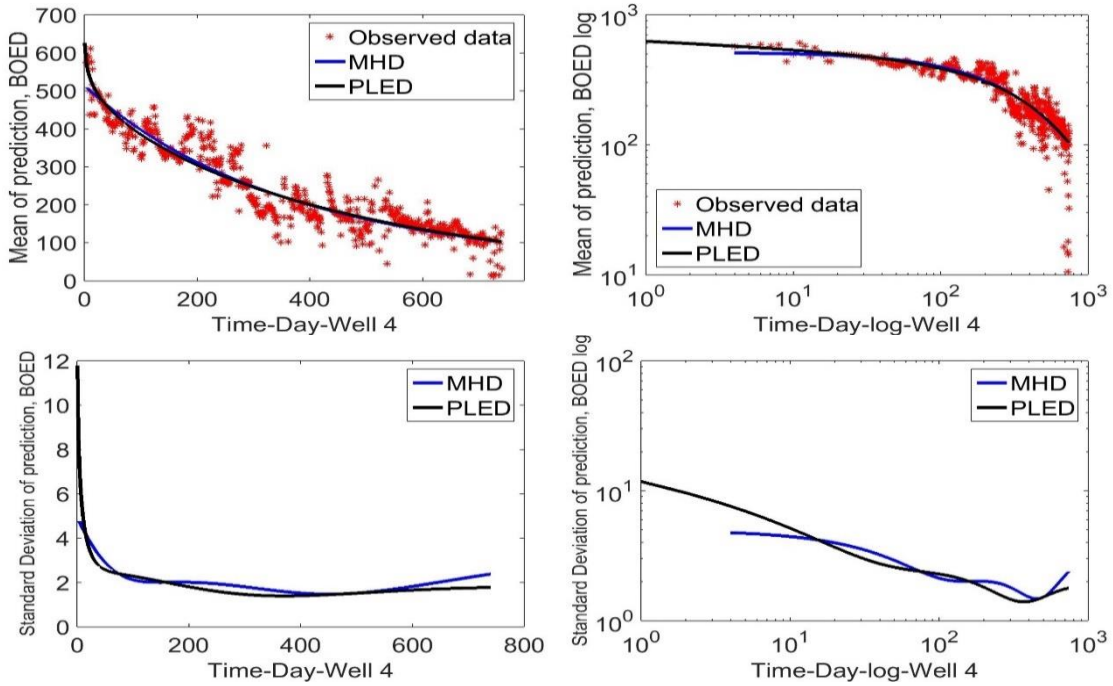

Fig. 4.34. Mean and standard deviation of MHD and PLED in normal and log-log scaled, Well 4

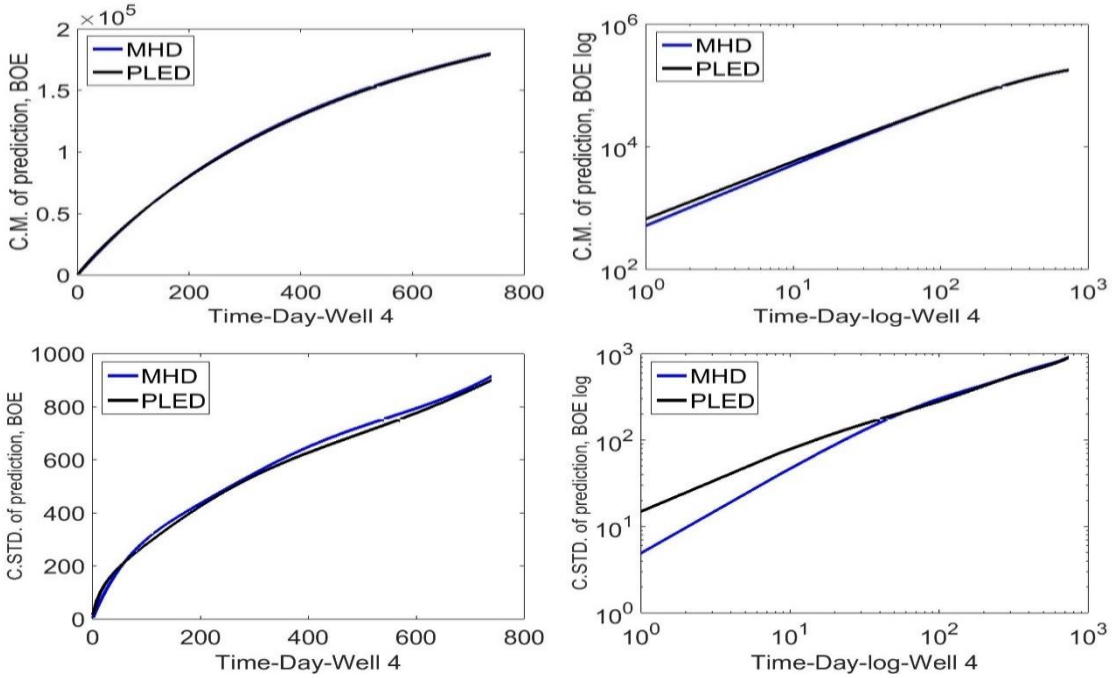

Fig. 4.35. Mean and standard deviation of the cumulative production of MHD and PLED in normal and log-log scales, Well 4

30 years' production time (daily basis)

Whereas the plot of mean of 30 years' production demonstrates the identical inclination, the degree of confidence in the result of MHD is relatively less than PLED.

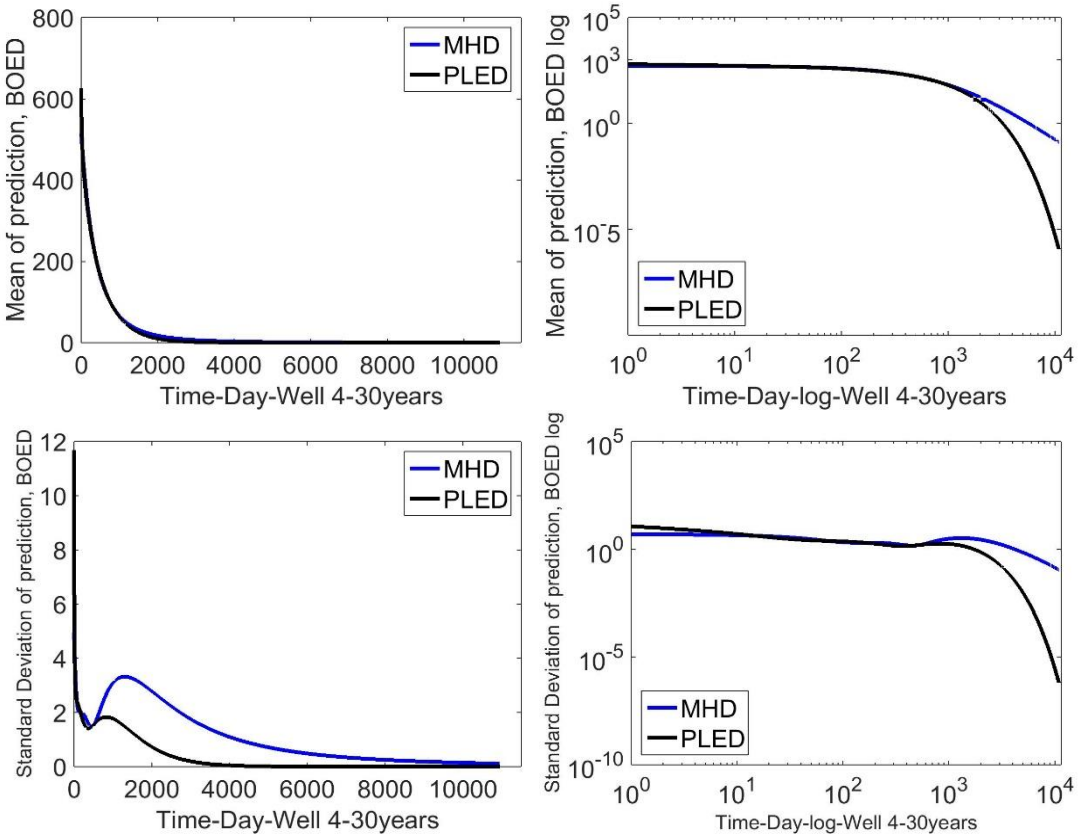

Fig. 4.36. Mean and standard deviation of 30 years production for MHD and PLED in normal and log-log scales, Well 4

30 years' production time (Cumulative)

Despite MHD provides more cumulative production during 30 years, the standard deviation of realization significantly sours in compare to the PLED.

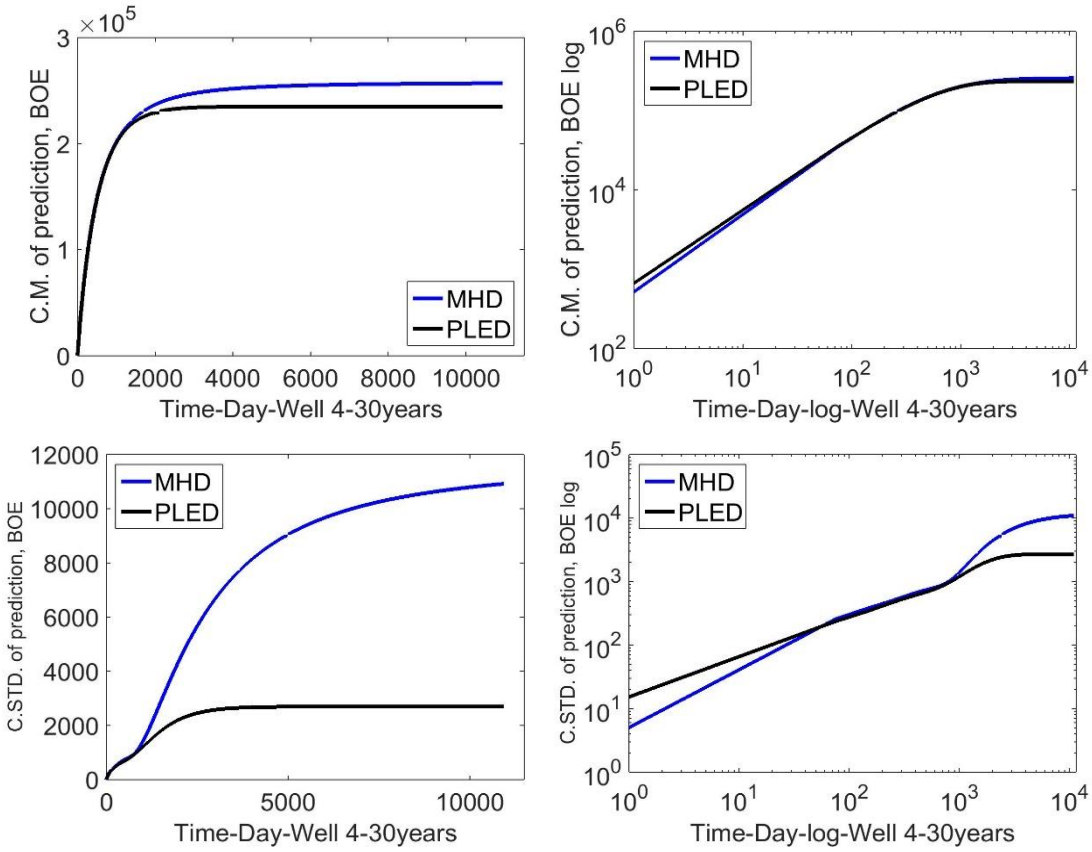

Fig. 4.37. Mean and standard deviation of 30 years cumulative production for MHD and PLED in normal and log-log scales, Well 4

The brief conclusion of this section can be found in section 8.1.3.

5. UNCERTAINTY ASSESSMENT OF A SEMI-ANALYTICAL MODEL APPLYING BAYESIAN PARADIGM FOR UNCONVENTIONAL RESERVOIRS

5.1. Overview

Semi-analytical models (Thermo-Chemo-Geomechanical) serve to evaluate the estimated ultimate recovery (EUR) of conventional/unconventional reservoirs. These models exhibit high parameter dimensionality, for which it becomes relevant to assess the influence of the parameters uncertainty into the model predictions. The objective of this work is to apply the Bayesian probabilistic calibration methodology that can delineate the impact of the experimental observations (field production data), the model predictions (from the semi-analytical model), and the expert judgment, onto EUR predictions.

We exploited a semi-analytical model to account for the heterogeneous characteristics of fractured wells and its influence on a shale gas reservoir. Three key parameters are defined as control variables: Maximum Permeability (K^0), Threshold Permeability (K_D^*), and Skin Factor (s). Together, the Markov Chain Monte Carlo (MCMC) method and the Metropolis-Hasting (MH) rule allow for the sampling of the posterior distributions corresponding to each one of these calibration experiments. From these distributions, it is then possible to generate likely model realizations representing the well's expected production. First and second order statistics of these realizations are used to assess confidence levels of production estimates from short to long term.

The probabilistic calibration permits to assess the expected value and variance of the model parameters. That is, for each calibration experiment, not only a probability density function can be retrieved for each random parameter, but also, the correlation structure among the parameters. The posterior generated for each calibration experiment, allows computing the cumulative means of the model production realizations. The use of Bayesian inference to assess the performance of the proposed analytic model for a given well production data shows that this uncertainty varies as the model parameters are allowed to vary.

5.2. Introduction

The semi-analytical production models allow us to simulate the behavior of the conventional and recently unconventional reserves incorporating the chemo-hydro-geomechanical characteristics of a play (Mattar and Anderson 2003; Fuentes-cruz, Gildin, and Valkó 2013; Tarrahi, Gonzales, and Gildin 2014). These models, however, in particular suffer from the high dimensionality of parameter space, which requires several in-situ or laboratory tests to obtain appropriate values and yet for some of them it is impossible to suggest a straightforward acquisition method. Thus, a considerable degree of uncertainty in the computation and model parameters' estimation is inevitable. The uncertainty in the observed data appears as another factor that also should be addressed. The inherent uncertainty may trigger the divergence of the outcome by overestimating or underestimating the estimated ultimate recovery (EUR). Therefore, it has become pertinent to assess the influence of uncertainty associated with the model parameter exerting probabilistic methods.

Bayesian probabilistic inference has become a reliable framework to assess the uncertainty embedded in the model results (Rotondi et al. 2006; Hoff 2009; Gelman et al. 2013; Bedi and Harrison 2013). In order to assess the Bayesian inference, it often become necessary to apply the Markov Chain Monte Carlo (MCMC) (Faming, Chuanhai, and Raymond 2010) and Metropolis-Hastings algorithm together to be able to feature the posterior. However, it should be taken into consideration that sometimes the complexity of the problem makes the MCMC algorithm computationally inefficient. Hence, these problems unavoidably lead the inverse solution to the more sophisticated Bayesian techniques such as Adaptive MCMC ensemble or Parallel Tempering frameworks which noticeably enhance the efficiency of sampling.

Roggero and Guérillot (1996) demonstrated the application of Bayesian analysis for the reservoir numerical modeling considering the uncertainty of geological parameters. A comparative study also derived to quantify the uncertainty of production forecasts in Floris et al. (2001). Furthermore, the application of deterministic and Bayesian probabilistic methods considering the scaling properties are broadly discussed in Vega, Rojas, and Datta-Gupta (2004). The significance of Bayesian analysis initially presented for the uncertainty assessment of Decline Curve models (Jimenez et al. 2005; Gong et al. 2011; Moridis et al. 2017).

Ibegbuna et al. (2012) employed the geometry of reservoir as the probabilistic random field to assess the behavior of Barnett Shale. A variety of techniques for forecasting a well production relying on the physical properties of a reservoir and corresponding limitations are provided and discussed in Lee and Sidle (2010). Moreover,

authors emphasized the importance of statistical approaches in analyzing production forecasting via comparison of most recent practical methods.

In addition, Zhang and Srinivasan (2005) exhibited the application of Markov Chain Monte Carlo (MCMC) and Metropolis-Hastings sampler (MH) in identifying the variation of permeability in reservoirs.

5.3. Methodology

5.3.1. Semi-analytical model

The exerted semi-analytical model incorporates the linear and exponential function planes alongside the dimensionless threshold permeability to address the non-uniform behavior of the permeability-area product (Fuentes-cruz, Gildin, and Valkó 2013). Besides, the model constructs the flow rate (Mscf/D) of production relying on several assembled Stimulated Reservoir Volume (SRV) to create a hydro-fractured block (Fig. 5.1). The semi-analytical model employs the physical, fluid and mechanical characteristics of the reserve with the contribution of sixteen parameters. Although the majority of contributed parameters could be invariably acquired applying field or laboratory tests, yet there is no meticulous approach to approximate the initial value of some of them. Thus, it is recommended to elaborate the probabilistic inversion solution to retrieve the corresponding statistics and assess the associated uncertainties.

In this study, the maximum permeability in millidarcy (K^0), threshold permeability (K_D^*), and skin factor (s) all together represent three key variables which serve as the

random field of the Bayesian paradigm. The semi-analytical model, hereafter, calls "forward model" that is a common expression in the Bayesian literature.

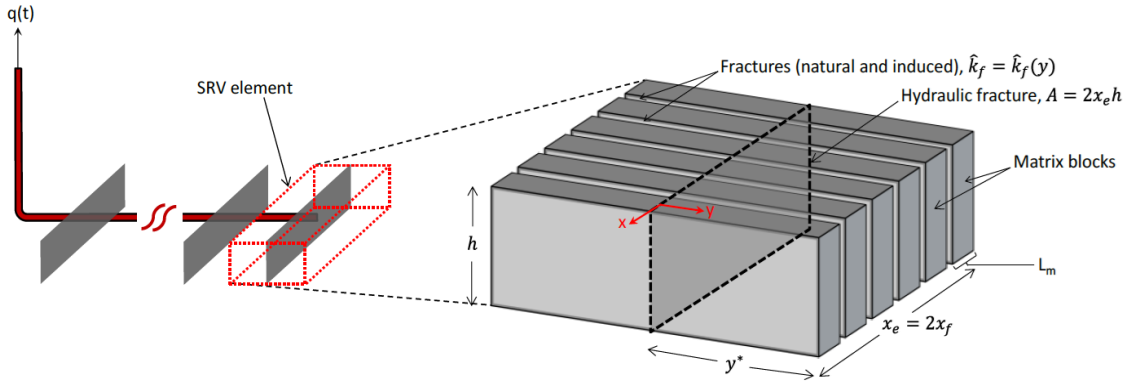

Fig. 5.1. The schematic configuration of the stimulated reservoir volume (SRV) and underlined production matrix, (reprinted with the permision from Fuentes-Cruz, Gildin, and Valkó 2014)

5.3.2. Bayesian and Markov Chain Monte Carlo (MCMC)

The applied method of Bayesian, MCMC are thoroughly discussed in section 2.3.

5.4. Experimental design

Table 5.1 illustrates the permutation of three random variables defining the case studies. Assimilation of variables eventually provides an insight, not only about the first and second order statistics of random field, but also the influence of the correlation structure on delineation of posteriori.

Table 5.1 The permutation of random variables

Parameters	One parameter	Two parameters	Three parameters
Maximum Permeability (K^0)	C_{K^0}	$C_{K^0-K_D^*}$	$C_{K^0-K_D^*-s}$
Threshold Permeability (K_D^*)	$C_{K_D^*}$	C_{K^0-s}	
Skin Factor (s)	C_s	$C_{K_D^*-s}$	

5.5. Observed data and parameter initial quantities

A set of Gas depletion log retrieved from the Barnett Shale development, provides the analysis observed data. The well log encompasses the total number of 988 days of production, which is also enclosed the entire required test results.

In addition, the initial values assigned to forward model are given in Table 5.2. The initial values, represent the state of expert believe about the parameters, thus we invoke them as the "expert belief."

Table 5.2 The semi-analytical forward model parameters and initial assigned values

	Parameter	Nomenclature	Initial values
1	Effective hydraulic- fracture half-length, ft.	x_f	400
2	Dimensionless threshold permeability	k_D*	0.0021
3	Maximum induced permeability, md	k^0	1.93E-03
4	Skin factor, -	s	0.216
5	Bottomhole flowing pressure, psia	p_{wf}	500
6	Half-length of SRV element, ft.	y*	552
7	Porosity, fraction	f_i	0.048
8	Rock compressibility, psi^{-1}	c_f	0
9	Water saturation, fraction	S_w	0.169
10	Formation thickness, ft	h	306
11	Reservoir temperature, °R		633.5
12	Initial pressure, psia	p_i	3115
13	Number of main hydraulic-fracture planes,	n_{HF}	0
14	Langmuir storage capacity, scf/ton	V_m	0
15	Matrix density, gm/cc	r_B	2.38
16	Langmuir pressure, psi	p_L	650

Fig. 5.2 depicts the observed data (Red Cross) coupled with the expert model (Blue solid line) in normal and log-log scales. Both plots exhibit a subtle match amid observation and expert model.

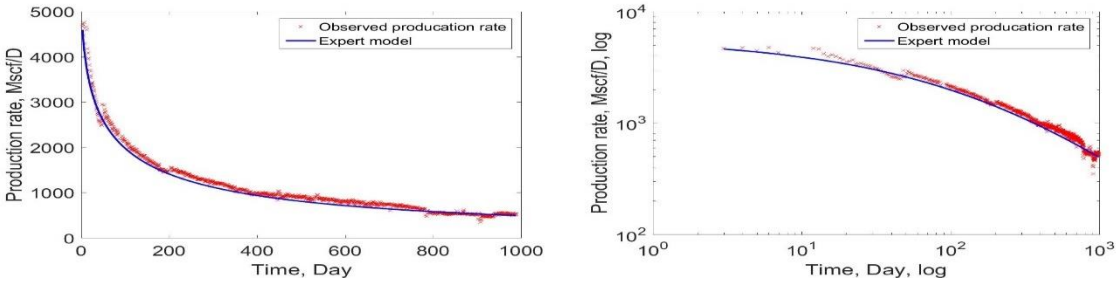

Fig. 5.2. The well production of observed data (red cross) and expert belief (blue solid line), left daily normal and right, log-log scales

Note that, not all set of observed data precisely demonstrates the cogent trend of production due to the outliers, hence requires to be filtered. In section 0, we featured a systematic filtering method that will become especially useful when a noticeable number of wells should be analyzed and applying a systematic filtering mechanism is inevitable.

5.5.1. Implementation of MCMC

In this section, we outline the methodology of MCMC in the format of implementation sequence of case $C_{K^0-K_D^*}$. Note that, we later demonstrate the results of entire case studies in the subsequent section.

5.5.1.1. Optimization

Optimization provides two major advantages featuring MCMC; comprising the initial values of random variables, and the possible shape of the likelihood distribution. The MCMC initial guess is achievable directly through the optimization when whole parameters but the random variables set as constant. Moreover, the approximation of likelihood distribution is developed from the residual (error) between the optimized fit curvature and the observed data. In this study, to address the optimization issue, the

method of non-linear least square method (NLS) is exerted. The fit of observed data and NLS optimization in normal and log-log scales is depicted in Fig. 5.3.

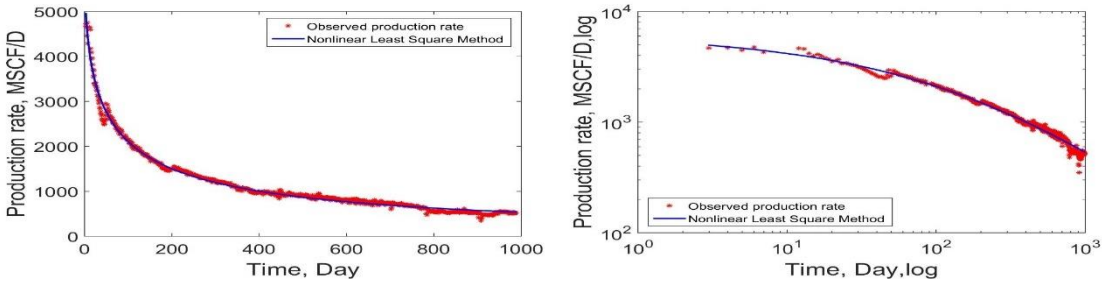

Fig. 5.3. A fit of NLS optimization and observed data, left, normal scale, and right, log-log scale

Fig. 5.4 , furthermore, shows the relative frequency histogram along with the cumulative density function plot of residuals.

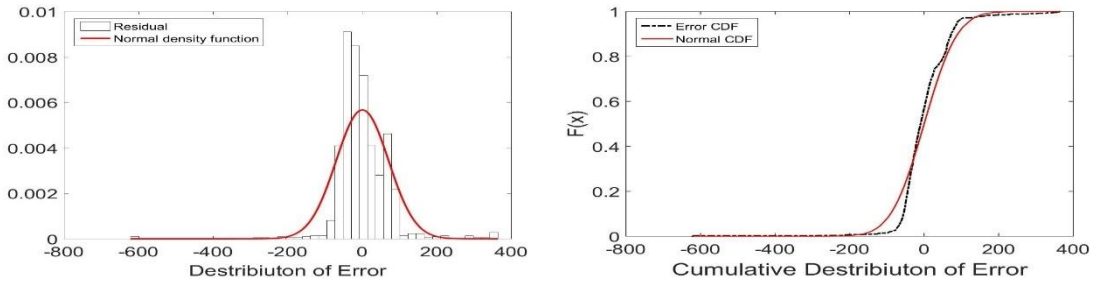

Fig. 5.4. Left, normal density function fits to the relative frequency histogram of residual, and left, the cumulative density function of residual and normal distribution

The appropriate synchrony amid the Normal distribution and residual in the above figure, also appreciates the initial perception of the likelihood distribution configuration. In addition, the standard deviation retrieved from the Normal density function can be later induced in the likelihood distribution of MCMC experiment.

5.5.1.2. MCMC experiment

We employed Eq.(2.3) and Eq.(2.6) to delineate the MCMC experiment. In this case, parameters K^0 and K_D^* serve as the random variables whilst the other parameters are kept constant. An important question considering the MCMC is how to detect the convergence of sampling. There are several methods available which among them the cumulative mean and standard deviation plots of MCMC experiment are more favorable. The MCMC is converged when both the cumulative mean and standard deviation plots become straight lines. The similar conclusion also can be drawn by eyeballing the MCMC experiment plots. The point which after that the cumulative plots become straight indicates the stationary condition or convergence of samples and is invoked as the "Burn-In point."

Fig. 5.5 depicts the MCMC experiment driven for 5e5 iterations on left and the cumulative mean and standard deviation of experiment on right. Moreover, the Burn-In point sets as 2e5.

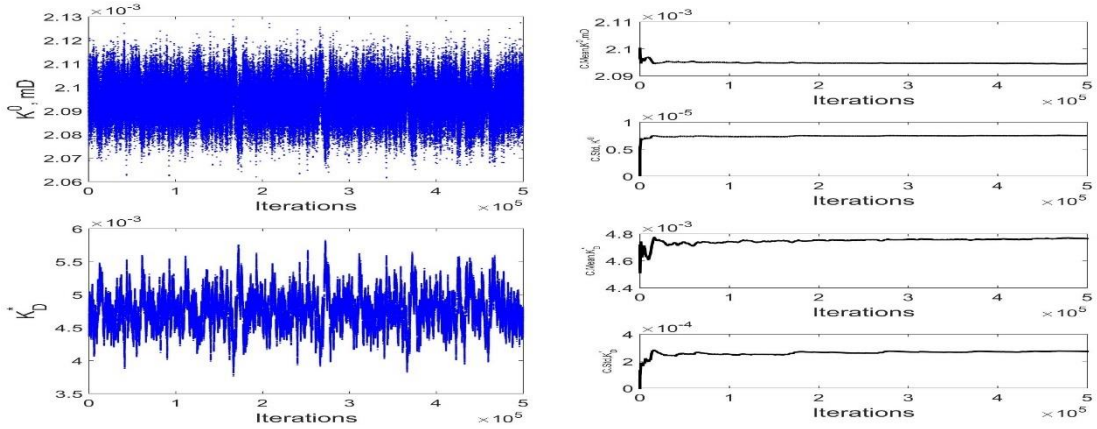

Fig. 5.5. Left, the MCMC experiment of K^0 and K_D^*, and right, the cumulative mean and standard deviation of experiments

5.5.1.3. Posterior distribution and statistics

The samples provided from MCMC, eventually generates the vector of posterior distribution that permits to capture the first and second order of statistics coupled with the correlation structure among the random variables. The associated plots presented in Fig. 5.6 provide a vivid insight regarding the possible distribution of random variables as well as the state of correlation among them. In this case, the analogy between the Normal distribution and drawn histograms is irrefutable. Besides, the plot of joint distribution exhibits the negative correlation between variables.

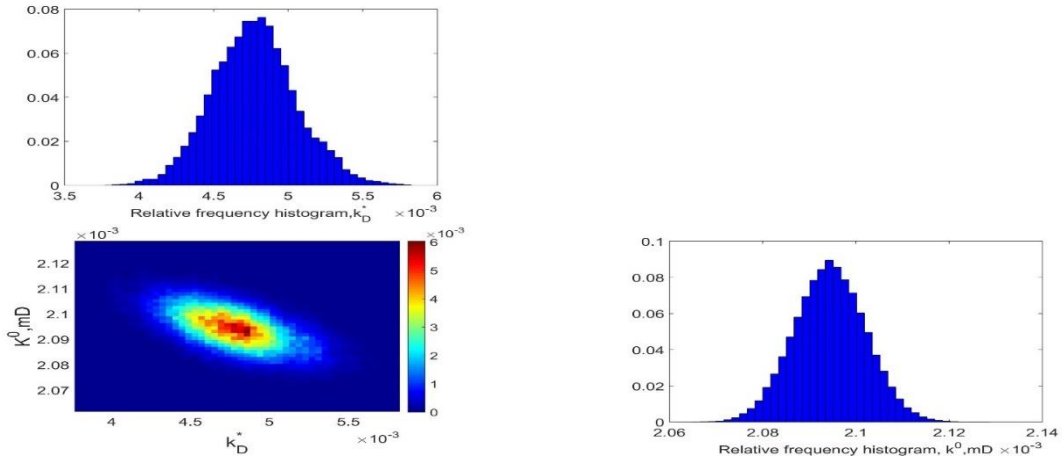

Fig. 5.6. The relative frequency histogram of random variables, left top and right down and the joint distribution of variables, left down

Table 5.3 illustrates the numerical quantities associated with the mean, standard deviation, mode and correlation coefficient of the vector of random field.

Table 5.3 The statistics of the vector of posterior and correlation coefficient of random variables

Parameters	Mean	Mode	Standard Deviation	Correlation coefficient
K^0, mD	2.094e-03	2.093e-03	7.621e-06	-5.829e-01
K_D^*	4.774e-03	5.011e-03	2.817e-04	

5.5.1.4. Current production realizations

Drawing a number of random samples (for instance 1000) from the posterior and plugging them into the forward model allows us to capture the most likely vector of together expect values and standard deviations of the model realizations. The vector of expected values then approximates the trend of observed data regarding the current production time. Then, the plot of standard deviation quantifies the level of confidence about the Bayesian analysis which also sometimes recalls as the uncertainty assessment. The plot of 1000 realizations along with the observed data, mean and standard deviation of realizations in both normal and log-log scales are demonstrated in Fig. 5.7.

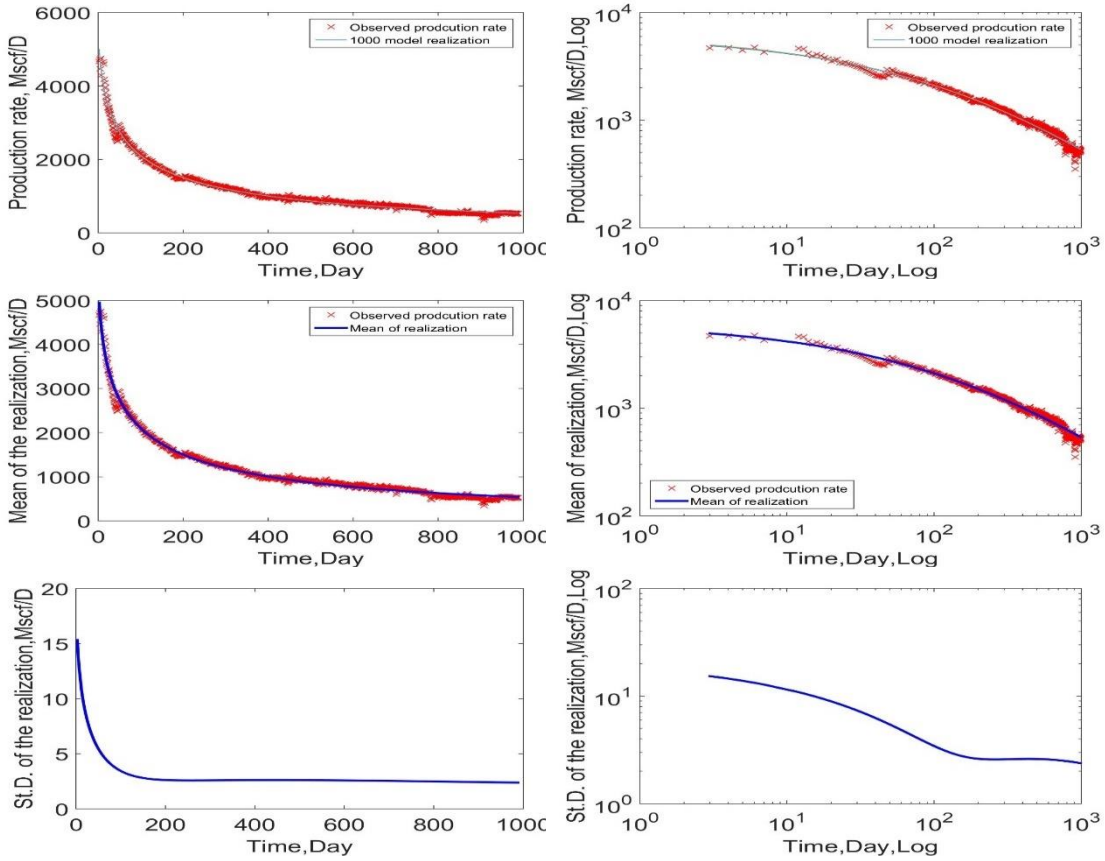

Fig. 5.7. Top, 1000 realization and observed data, middle, mean and down, standard deviation of realization in left, normal and right, log-log scales

The plot of standard deviation depicts the state of certainty by progressing in time. Despite the certainty is less at the beginning of the process, the level of confidence sours after 200 days of depletion.

5.5.1.5. 20 years' production realization

A substantial application of the Bayesian analysis appears as a mean to anticipate the well production for long term. The extrapolation of realizations has become a reliable practice to approximate the well future depletion. Fig. 5.8 presents the 1000 realization plot together with the mean and standard deviation or realization in both normal and log-log scales.

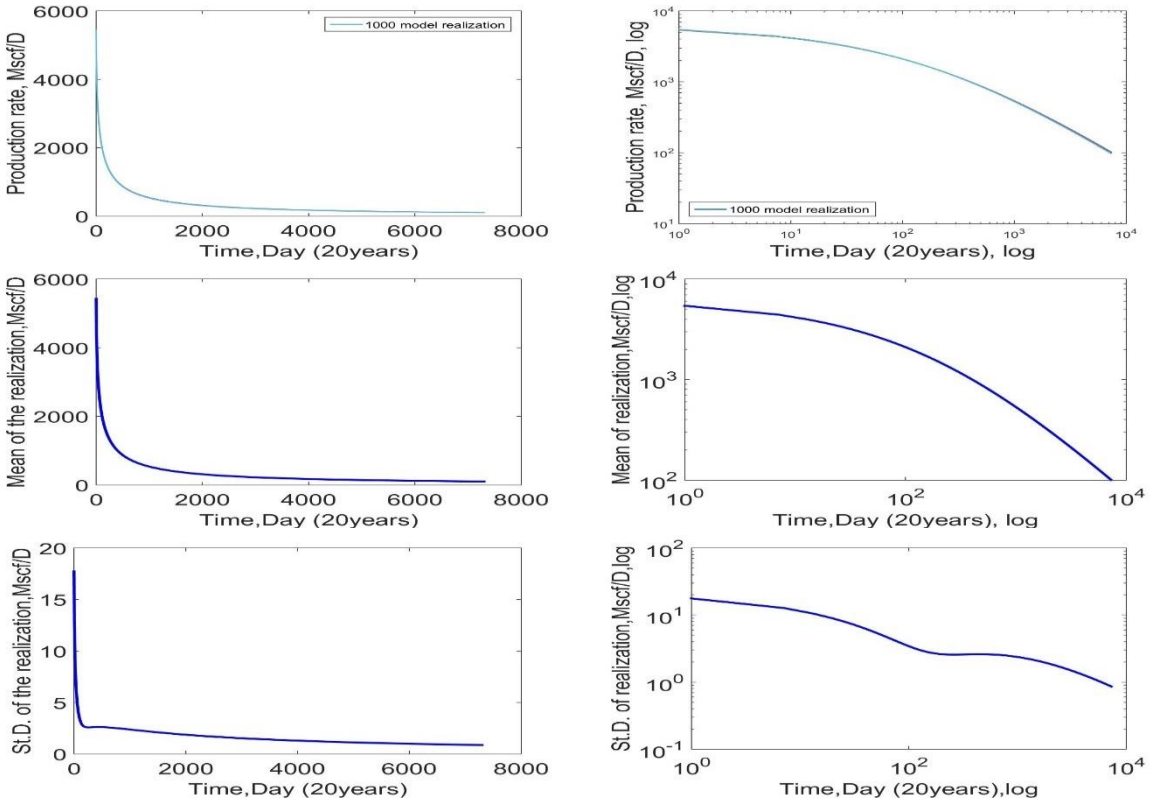

Fig. 5.8. Top, 20 years' realization of production, middle, mean and down, standard deviation of 20 years' realization in left, normal and right, log-log scales

Standard deviation plot in Fig. 5.8 provides the significant improvement in the level of confidence after 2000 days. This plot allows us to perceive the behavior of well production for a long term and the necessity of acquiring data merely up to 2000 days.

5.6. Results and discussion

The analogy between statistics of entire case studies are illustrated in Table 5.4. According to Table 5.4, the expected value and mode of parameter K^0 have not been noticeably altered, however, the level of confidence is dropped from one parameter to three parameters cases.

Table 5.4 Mean, mode and standard deviation retrieved from the experimental designs

Parameter	Inference	Case studies						
		C_{K^0}	$C_{K_D^*}$	C_s	$C_{K^0-K_D^*}$	C_{K^0-s}	$C_{K_D^*-s}$	$C_{K^0-K_D^*-s}$
K^0, mD	Mean	2.15e-03	-	-	2.09e-03	2.35e-03	-	2.18e-03
	Mode	2.15e-03	-	-	2.09e-03	2.35e-03	-	2.09e-03
	Standard Deviation	6.76e-06	-	-	7.62e-06	1.77e-05	-	4.28e-05
K_D^*	Mean	-	1.04e-02	-	4.77e-03	-	8.37e-03	3.69e-03
	Mode	-	9.42e-03	-	5.01e-03	-	8.02e-03	4.96e-03
	Standard Deviation	-	5.83e-04	-	2.82e-04	-	4.56e-04	5.13e-04
s	Mean	-	-	1.89e-01	-	2.52e-01	1.93e-01	2.29e-01
	Mode	-	-	1.89e-01	-	2.55e-01	1.93e-01	2.17e-01
	Standard Deviation	-	-	1.09e-03	-	3.09e-03	1.12e-03	6.23e-03

Also, it is observed that the statistics of K_D^* varies from case to case, nevertheless, the range of variation is almost similar. Despite the mean and mode of parameter s exhibits a

fluctuation, case by case, the certainty diminished for $C_{K^0-K_D^*-s}$. The general pattern which can be traced in all cases is that by adding more random variables the level of confidence decreases, but the reduction rate is depended to each individual case study.

Moreover, Table 5.5 provides the state of correlation coefficient among random variables in two different scenarios, two parameters together and three parameters together. Considering the corresponding quantities, it is observed that by increasing the size of random space, not only the order of correlation coefficient is changed, but the sign of them is also influenced.

Table 5.5 The correlation coefficient between random variables

Case studies	$K^0 - K_D^*$	$K^0 - s$	$K_D^* - s$
Two parameters ($C_{,-,}$)	-0.583	0.911	0.199
Three parameters ($C_{,-,-,}$)	-0.929	0.983	-0.9013

For instance, the magnitude of correlation coefficient among parameters $K_D^* - s$ indicates that the state of correlation has substantially variates from the less positively correlated in two parameters case to the highly negative correlated in the three parameters case.

Following, several comparison plots drawn from the current and 20 years' realizations. The current time mean realizations of whole experimental designs together exhibit a relatively identical match to the observed data, whilst the log-log scaled reveals a slight divergence in the expert model (Green solid line, Fig. 5.9, top). Observing the same plot, the 20 years' realization, and the expert model underestimates the long-term production, however, entire study cases produce similar curvatures.

In addition, the standard deviation of experimental designs in Fig. 5.10 introduces more heterogeneity in the results. Despite beginning all current production time plots at a point about 30 Mscf/D, the standard deviation plot of K_D^* presents a completely reversed behavior and starts in proximity of zero Mscf/D. Also, the plot of skin factor in log-log scale depicts a sour in the level of confidence after a drop around 400 days. Generally, the state of certainty of analysis becomes vary from case to case.

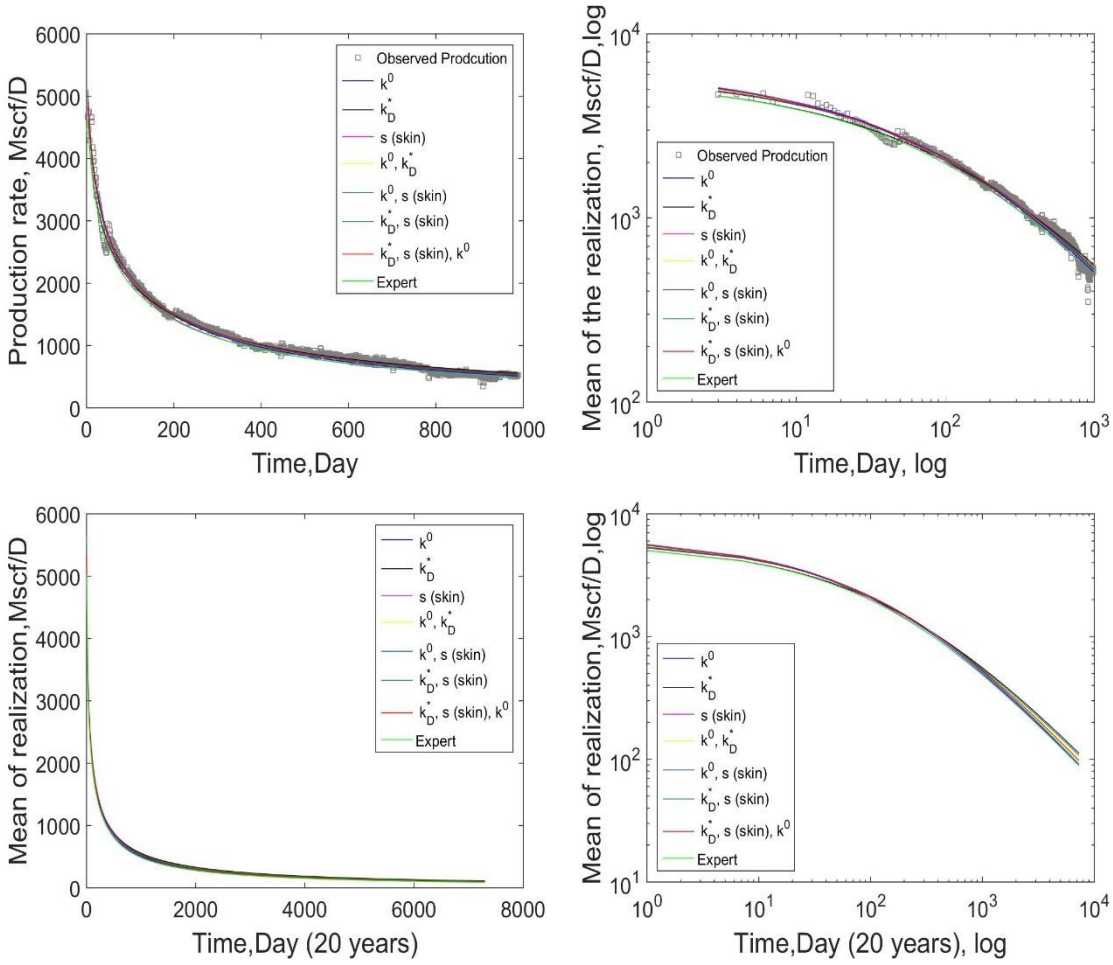

Fig. 5.9. Analogy between the mean of realization of case studies and expert model for the short and long term of production

Furthermore, a similar analogy can be seen in the 20 years' production. Note that, regardless of dissemination in the outcomes of standard deviation of realizations, their order of magnitude in comparison to the real scale of production is negligible.

An extra feature to demonstrate the well depletion is the form of cumulative production (MMscf). Considering the produced plots, it is evident that, the cumulative plot in some cases develops a better perception about the production rate.

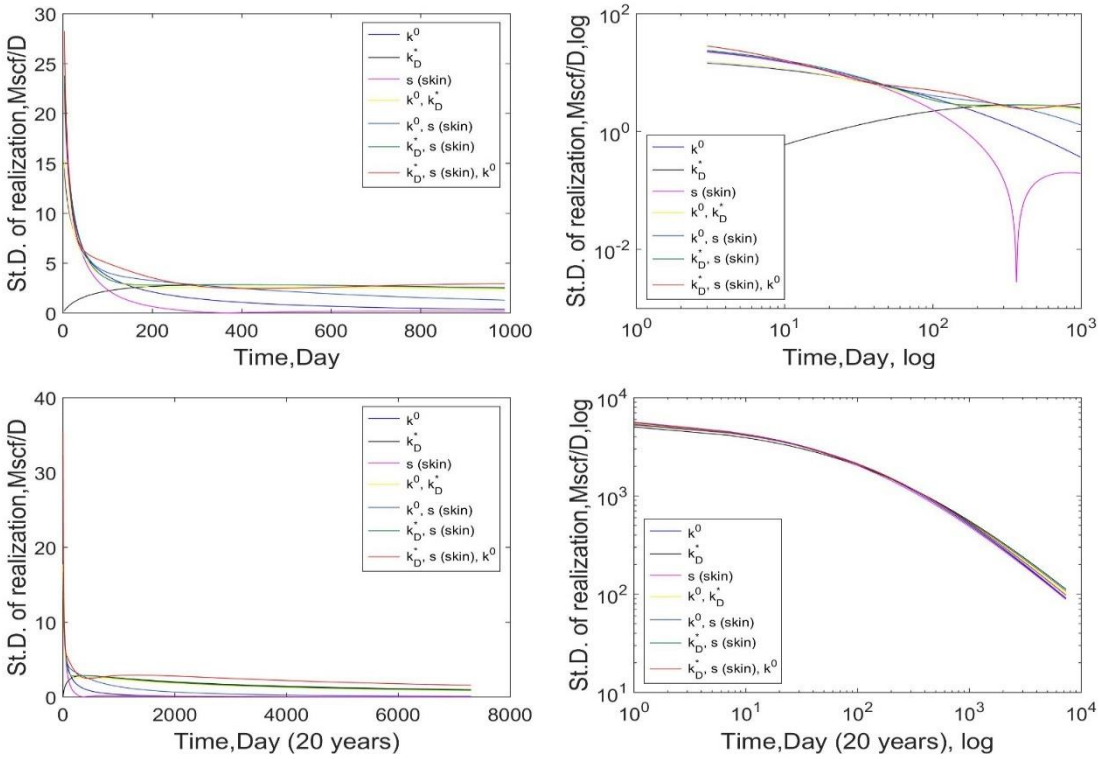

Fig. 5.10. The standard deviation of realization of whole experimental designs regarding the current and long-term production

The current and 20 years' production time of cumulative mean plots (Fig. 5.11) depict a set of distinguished curvatures by progressing in time. The expert model and then case C_s exhibit the lower cumulative production rate both in the current and 20 years plots.

In analogy with the cumulative mean, the cumulative standard deviation is also can be drawn to elucidate the state of confidence of Bayesian analysis. From Fig. 5.12 it is alluded that, C_s provides less uncertainty in both the current and 20 years' production.

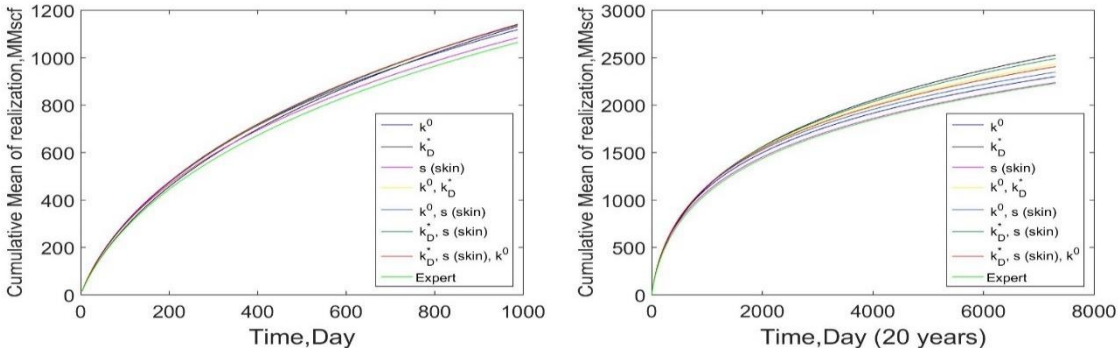

Fig. 5.11. The ensemble of cumulative production rate, left, current and, right, 20 years' production time

Furthermore, it is recognized that, although the case $C_{K_D^*-s}$ shows less confidence for the current time, the case $C_{K^0-K_D^*-s}$ erratically demonstrates a great degree of uncertainty regarding the long-term production.

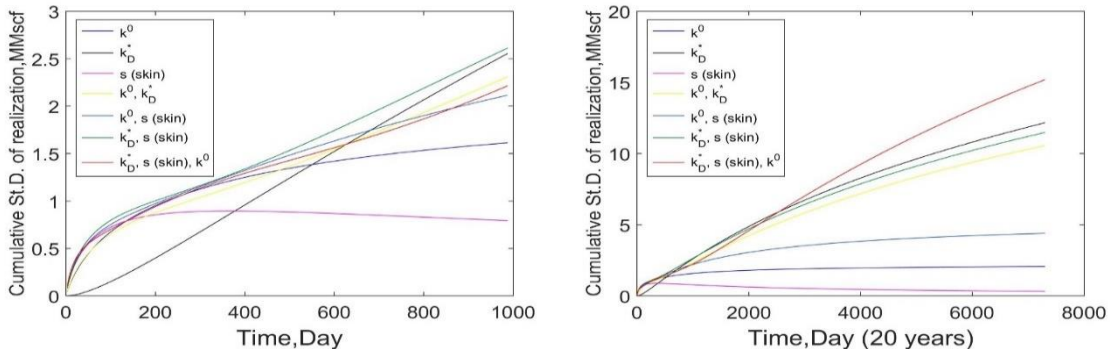

Fig. 5.12. The cumulative standard deviation of whole cases, left, the current, and right, the 20 years' production

The corresponding conclusions are provided in section 8.1.4.

6. APPLICATION OF THE SEQUENTIAL BAYESIAN UPDATING ON THE UNCERTAINTY QUANTIFICATION FOR AN EMPIRICAL OIL AND GAS PRODUCTION MODEL

6.1. Overview

Despite the recent substantial developments in the methods that deal with the Bayesian analysis, yet a straightforward procedure requires to be elaborated to feature the implementation mechanism of adding a new set of observations to the already existed data. Which is for, it is pertinent to investigate the significance of updating the state of belief by exerting the posterior of previously performed Bayesian paradigm as the prior of the current step, in compare to the merely derived Bayesian inference based on the aggregate vector of previous and updated observations. We also applied the Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) framework to develop the Bayesian inference coupled with a new algorithm to construct a hybrid prior that allows incorporating the posterior of the previous Bayesian assessment. Additionally, whilst four sets of well depletion logs, retrieved from the Eagle Fort Shale, constitute the observed data; the Modified Hyperbolic Decline (MHD) Curve model delineates the well Estimated Ultimate Recovery (EUR). The performance evaluation of the Bayesian regular and updating approaches, subsequently implies the persistent reduction in the attained uncertainty using the updating technique for the short and long term of production.

6.2. Introduction

The Bayesian paradigm has recently become popular amidst scholars and practitioners, due to its definite advantages over other probabilistic inference techniques. Also, Bayesian filtering (Z. H. E. Chen 2003; Lauritzen 2008; Sarkka 2013) denotes a general term for updating of the Bayesian inference when the new data continuously becomes available and encompasses several methods such as Sequential Monte Carlo updating (Doucet, Godsill, and Andrieu 2000), Particle filtering (D. S. Lee and Chia 2002; Arulampalam et al. 2002; Carvalho et al. 2010; Andrieu, Doucet, and Holenstein 2010; Lopes et al. 2012) or in particular, Kalman filtering (R. Chen and Liu 2000) . However, Bayesian filtering technically is pertinent to the case of closed-form problems while for the class of non-closed-form problems more sophisticated methods are required (Lauritzen 2008).

Andrieu, Freitas, and Doucet (1999), outlined the application of MCMC in the online updating in association with the Importance Sampling and Reversible Jump MCMC methods. The method is mainly useful when a closed-form model is inaccessible or unknown. More scholars have gradually incorporated the evolution of MCMC in the sequential Bayesian updating (Andrieu, Doucet, and Holenstein 2010; Y. Yang and Dunson 2013; Septier and Peters 2016). Despite all aforementioned constructed methodologies which deal with the stochastically updating data when the parameter space size is either known or unknown and some perception regarding the target distribution coexisted, the issue of selection an appropriate prior is yet to be addressed. In the case of closed-form problems it is plausible to derive Bayesian equations mathematically to

retrieve the Bayesian prior and posterior predictive distributions to enable utilizing them in the Bayesian updating analysis (Gelman, Meng, and Stern 1996; Held, Schrodle, and Rue 2010).

The non-informative prior distribution is hypothesized whereas there is less knowledge about the possible shape of the target distribution statistics and that, the random samples are assumed to be independent and identically distributed (i.i.d). In such a case, applying Markov Chain Monte Carlo (MCMC) is inevitable (Gelman et al. 2013). MCMC along with the Metropolis-Hastings (MH) (Metropolis et al. 1953; Hasting 1970) algorithm integrates random samples over the entire parameter space to eventually converge to the true posterior distribution. MH algorithm renders an acceptance-rejection criterion by drawing random samples from a proposal distribution when the target distribution is inaccessible.

The conjecture of a well's depletion applying the production models has become a key tool to assess the well Estimated Ultimate Recovery (EUR) in either short or long period of time for Hydrocarbon plays (Robertson 1988; Fetkovich, Fetkovich, and Fetkovich 1996). In addition, dealing with the continuously incoming data raised the concern of how to absorb the outcomes of former analysis with the new data to enhance the resolution of future modeling. Considering the wide application of after Arps empirical models (Arps 1944) due to their readily implementation and few constitutive parameters, the Modified Hyperbolic Decline curve model (MHD) is selected to serve as a reliable framework to reproduce the well behavior (Robertson 1988; Ilk et al. 2008). MHD model hereafter is referred as the "Forward model" in the Bayesian analysis.

We established a mechanism that by obtaining the vector of posterior distribution from the former MCMC sampling and updating the observed data after a time interval, one shall be able to assimilate the current information as the prior of the next step of the MCMC experiment. In order to implement the sequential MCMC updating, firstly a sampler is required to readily and precisely constitute MCMC. Therefore, a novel methodology which is evoked as the Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) is exerted (section 2.3.2). PSAMH by contributing parallelized concurrent chains automatically adapts the step size of the proposal distribution in the Metropolis-Hastings algorithm to optimize the acceptance rate. PSAMH is initially originated from the Adaptive MCMC ensemble methods which have been widely iterated in recent literatures (Haario, Saksman, and Tamminen 2001; Bedard and Rosenthal 2008; Andrieu and Thoms 2008; Gareth O. Roberts and Rosenthal 2009; Rosenthal 2011). The applications of PSAMH exploiting empirical decline curves and semi-analytical models are thoroughly discussed in section 4, Moridis et al. (2017) and section 5, respectively.

Secondly, we constructed a hybrid prior which encompasses the relative frequency histogram of the retrieved posterior along with the non-informative distribution when it becomes necessary which together serve as the next step prior and will discuss thoroughly later.

The results attained from the contribution of aforementioned hybrid mechanism and PSAMH algorithm, incorporate to draw the forward model realizations that eventually appear as the mean and standard deviation of process. Whilst the mean of realization indicates the general trend of well behavior, the standard deviation regulates the inherent

uncertainty in the analysis. The comparison between the standard deviation of realizations retrieved from the updating and regular techniques utilized to denote the efficiency of updating method.

6.3. Methodology

6.3.1. Modified Hyperbolic Decline curve

We exerted the Modified Hyperbolic Decline (MHD) in this section which previously discussed in section 4.3.1.

6.3.2. Markov Chain Monte Carlo (MCMC) and Parallel Scaled Adaptive Metropolis-Hastings (PSAMH)

The implemention sequence of MCMC and PSAMH framework are comprehensively elaborated in section 2.3.2.

6.3.3. Last time interval's posterior as the current prior

MCMC draws thousands of samples to gradually converge to the true posterior. Obtaining a set of posteriors, provides an opportunity to develop the prior of the next updating step. The multivariate relative frequency histogram is proposed as the key technique to impose the last vector of posterior as the current prior. However, often by updating the observed data, it becomes inevitable that the sampler searches the parameter space beyond the acquired posterior space. In order to address this issue, we impose the non-informative prior (Uniform distribution with a wide range of boundary values) as the enclosed term to switch to it when it becomes pertinent. Fig. 6.1 demonstrates a visual

configuration of a unimodal hybrid prior which is the combination of the last step posterior and the non-informative prior.

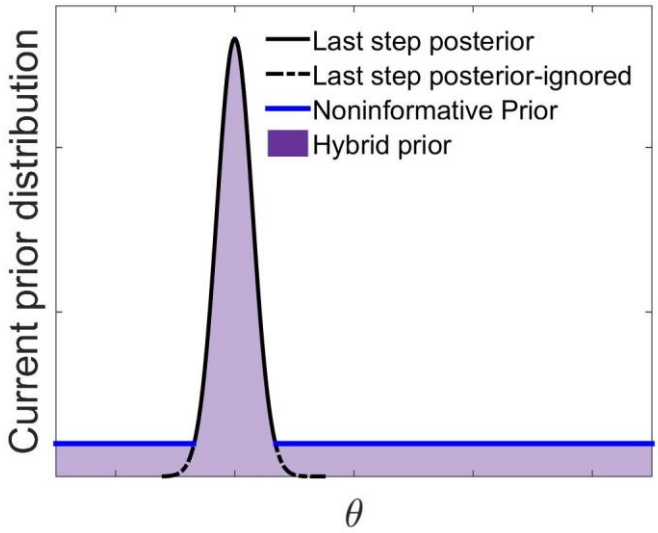

Fig. 6.1. Demonstration of combination of posterior and non-informative prior as a hybrid prior

6.3.4. Generating the current step prior

Succeeding, the pseudo-steps employed to constitute the hybrid prior density function is featured.

6.3.4.1. Determination of the bin size and range of histogram

The following two subsections develop a method that allows generating the histogram of a d dimensional posterior.

- Set the number of bins (n_{bin}) of the relative frequency histogram drawn from the vector of last step posterior.
- Applying the bin-size, split the range of vector of variables from the past posterior into n_{bin} segments.

$$\Delta\boldsymbol{\theta} = \frac{\boldsymbol{\theta}_{max} - \boldsymbol{\theta}_{min}}{n_{bin}} \qquad (\boldsymbol{\theta} = \theta_1, \theta_2, \dots, \theta_d) \tag{6.1}$$

Where, $\boldsymbol{\theta}_{max}$ and $\boldsymbol{\theta}_{min}$ denote the maximum and minimum values of parameters ($\boldsymbol{\theta}$) retrieved from the posterior.

Next, let

$$range(\boldsymbol{\theta}) = (\boldsymbol{\theta}_{min} + ii * \Delta\boldsymbol{\theta},\ \boldsymbol{\theta}_{max}) \qquad (ii = 0,1, \dots, n_{bin} - 1) \tag{6.2}$$

$range(\boldsymbol{\theta})$ permits to establish a d dimensional matrix with equal number of bins at each individual direction.

6.3.4.2. Generating the multivariate relative frequency histogram

- Reckon the number of set of variables that lays inside each bin ($range(\boldsymbol{\theta})$) and store them in a $M = {n_{bin}}^d$ matrix. Note that, the retrieved matrix encompasses several zero value components.
- Divide the components of matrix, *M* by the vector length of the last step posterior (n_{total}) to derive the relative frequency histogram, $\pi_{last\ step\ posterior}(\boldsymbol{\theta})$.
- Also, obtain the non-informative prior of random variables.

$$\pi_{non-informative}(\boldsymbol{\theta}) \sim U(\boldsymbol{\theta}_{lower-bound}, \boldsymbol{\theta}_{upper-bound}) \tag{6.3}$$

- Substitute all components of $\pi_{last\ step\ posterior}(\boldsymbol{\theta})$ which are smaller than $\pi_{non-informative}(\boldsymbol{\theta})$ with the computed non-informative probability values (Eq. (6.3)). Implementing the above steps, a hybrid prior matrix incorporating the past posterior and non-informative distribution is achieved, $\pi_{current-prior}$. Note that either limits (lower and upper limits) of the current hybrid prior become the non-informative prior boundaries that is more broaden. The generated hybrid prior is

a matrix developed an equal possibility for all set of samples inside or outside the range of the last step posterior when the probability of current prior appears less than the non-informative prior.

A typical configuration of a bi-dimensional histogram is depicted in Fig. 6.2 with the number of bins set as 5. Additionally, considering the plot, some extra information such as the either boundaries ($\boldsymbol{\theta}_{min}$, $\boldsymbol{\theta}_{max}$), ranges of variables, the rectangular volumes representative of the count of samples and the bin interval are pointed by arrows.

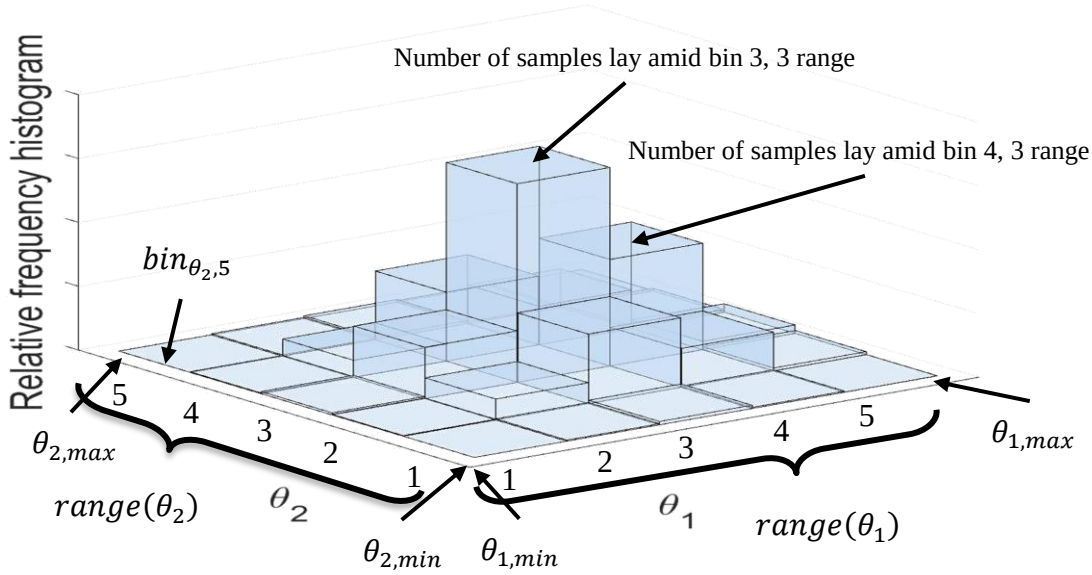

Fig. 6.2. A schematic configuration of relative frequency histogram and bin arrangements for a 2D random space

6.3.5. Imposing the hybrid prior distribution into the updating MCMC

- Draw a set of samples from the proposal distribution, $\boldsymbol{\theta}_i$.
- Determine the position of the new set of drawn samples associated with the $range(\boldsymbol{\theta})$ of the hybrid prior matrix. Indicate the corresponding probability

quantity from the prior matrix and set it as the current prior distribution value. Note that, if the range of current drawn samples is outside of the boundaries of $range(\boldsymbol{\theta})$, the $\pi_{non-informative}(\boldsymbol{\theta})$ is assigned as the associated current prior.

- Pursue the PSAMH sampling as instructed earlier.

6.4. Wells' production logs

Table 6.1provides four wells' production data (Well 1 to Well 4) from the Eagle Ford Shale which serve as the observed data. We delineated a systematic framework in section 0 to filter a raw depletion log to remove the unnecessary outliers.

Table 6.1 Wells' name and production duration before and after filtering

Well name	Well 1	Well 2	Well 3	Well 4
Duration –Raw (Day)	738	705	721	740
Duration –Filtered (Day)	665	592	609	650

Utilizing the same mechanism, the raw data of wells are filtered and demonstrated in Fig. 6.3. Note that, in order to save the space in this paper, we avoided repeating the filtering procedure here, and thus encourage readers to study the above invoked paper to learn how to implement the filtration routine.

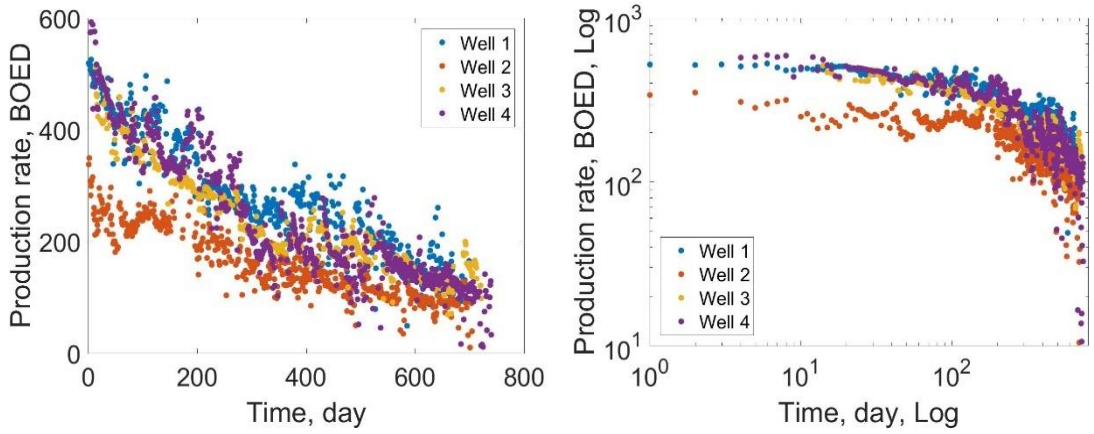

Fig. 6.3. Filtered wells' production data in the regular and log-log scales

In addition, the same procedure is employed regarding updating sequence from the first to the last time interval (Fig. 6.4, Well 1).

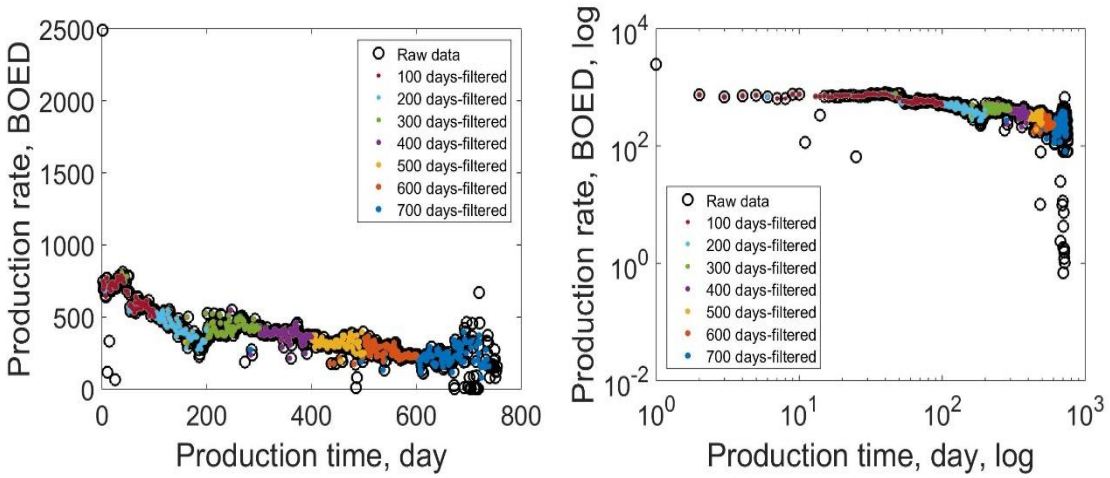

Fig. 6.4. Sequential updating filtered and raw data, left, the regular and right the log-log scales for Well 1

6.5. Experimental design and input data

The computational experiment integrates seven updating sequence with 100 days' intervals. Moreover, another test for the entire production time (Current) which is about 700 days for all four wells are derived and compared to the last updating results corresponding to each case (Table 6.2).

The procedure that features acquiring the MCMC initial values will be elaborated later. The range of random variables for the Uniform distribution set to b, [1e-4, 2], D_i, [1e-4, 20] and q_i (BOED), ([0.5, 2.5]×max(observed data)).

6.6. Implementation of the sequential MCMC updating

6.6.1. 100 days and the current production time

Table 6.2 Initial values of random parameters associated with each updating step

Name	Time (day) / Variable	100	200	300	400	500	600	700	Current
Well 1	b	1e-4	4.2e-3	4.9e-4	1.96	1.98	1.99	1.89	1.38
	D_i	3.7e-3	3.6e-3	4.3e-3	6.2e-3	5.1e-3	5.4e-3	5.4e-3	4.1e-3
	q_i (BOED)	772.6	771.08	791.11	796.45	764.78	775.94	782.65	757.16
	$\sigma_{likelihood}$	31.68	29.52	50.78	42.98	40.53	41.02	45.24	45.24
Well 2	b	2.0	0.79	0.36	7.5e-3	1.6e-3	6.3e-4	2.19e-4	1e-4
	D_i	2.6e-3	2.2e-3	2.56e-4	1.2e-3	1.7e-3	1.7e-3	1.7e-3	1.7e-3
	q_i (BOED)	270.14	268.27	307.39	266.16	279.23	279.24	279.15	277.50
	$\sigma_{likelihood}$	16.21	21.43	24.12	23.79	23.51	20.80	19.97	19.97
Well 3	b	2	1.95	9.14e-4	5.2e-4	4.77e-4	0.54	0.63	0.73
	D_i	1.3e-2	1.29e-2	2e-3	2e-3	2.3e-3	2.7e-3	2.8e-3	3e-3
	q_i (BOED)	623.47	619.43	435.87	437.22	452.2	461.15	462.19	475.95
	$\sigma_{likelihood}$	33.55	14.35	13.95	13.96	18.89	18.44	19.52	19.52
Well 4	b	2	1.93	1.91	1.96	3.3e-3	6.6e-4	1.2e-2	0.24
	D_i	1.2e-2	1.2e-2	3.5e-3	3.8e-3	2.5e-3	2.5e-3	2.3e-3	2.7e-3
	q_i (BOED)	626.54	625.25	507.78	516.27	516.45	514.7	500.3	514.39
	$\sigma_{likelihood}$	18.3	27.18	33.75	41.17	39.0	38.59	34.21	34.21

Bayesian analysis of the first 100 days as well as the current time production (about 700 days) corresponds to the regular PSAMH sampling which is comprehensively outlined in section 4. The concise vertices of pseudo-sequence are recapitulated as subsequent

- Construct the nonlinear least square optimization to capture the initial quantities of random variables coupled with the standard deviation of residual. Additionally, the histogram or cumulative density function plots of residuals permit to assess the possible configuration of the MCMC likelihood, $\sigma_{likelihood}$.
- Draw random samples exploiting the PSAMH framework.

Note that, the Normal distribution serves as the likelihood in the PSAMH sampling. PSAMH is iterated for 2e6 times to ensure attaining enough samples after achieving the stationary condition (Burn-In point).

- Plot the MCMC experience along with the cumulative mean and standard deviation of random field to identify the state of convergence and then set the Burn-In point.
- Assess the first and second order of statistics of posterior together with the correlation structure amid random variables.
- Draw the short and long-term production realizations by plugging the vector of posterior into the forward model.
- Plot the expected value and standard deviation of model realization for both the current and 30 years' depletion.

6.6.2. 200 to 700 days updating

Obtaining the posterior of the first time-interval (100 days), allows to induce the expected values of random variables as the next updating step initial values (See Table 6.2, for instance production time of 200 days). Furthermore, applying the hybrid prior distribution mechanism, the prior of the sequential updating becomes available. The rest of computation steps resemble to the PSAMH framework. The aforementioned sequence should be replicated to the last production time interval (700 days).

6.7. Results and discussion

In this section, we initially feature the implementation sequence of Well 1 and later demonstrate the corresponding comparison results obtained from all case studies.

Fig. 6.5 encompasses eight compartments, illustrates the evolution of three-dimensional joint distribution of model random variables for all updating sequences on left. In addition, the associated MCMC experiment scaled plot of variable b together with the current time extrapolation of the mean of realizations coupled with the observed data are demonstrated at right, top and down of each time interval. It is alluded that the order of magnitude of variable b is not constant and changes by progressing in time (Seshadri and Mattar 2010; Ilk et al. 2008), for which it becomes pertinent to evaluate the influence of updating method by evaluating this variable. Moreover, the mean of realization curvature indicates the overall perspective of prediction relying on different time intervals.

According to Fig. 6.5, posterior layout of 100 days provides a unimodal oval shape distribution whilst proceeding to 200 days, the volume manifests a multimodal arbitrary distribution with noticeably smaller variance. Simultaneously, examining the discrepancy of variable b in either configurations reveals a substantial declination in the 200 days' plot, which substantiates the associated shrinkage in the posterior space. However, the posterior of 300 days' production displays again a unimodal oval shape with slander variance. Following the sequence of plots from 400 to 500 days, demonstrates the enhancement in the confidence of analysis. Nevertheless, considering the posterior and MCMC experiment plots of 600 days interval, connotes a decline in the level of certainty which is relevant to the scattered nature of observed data.

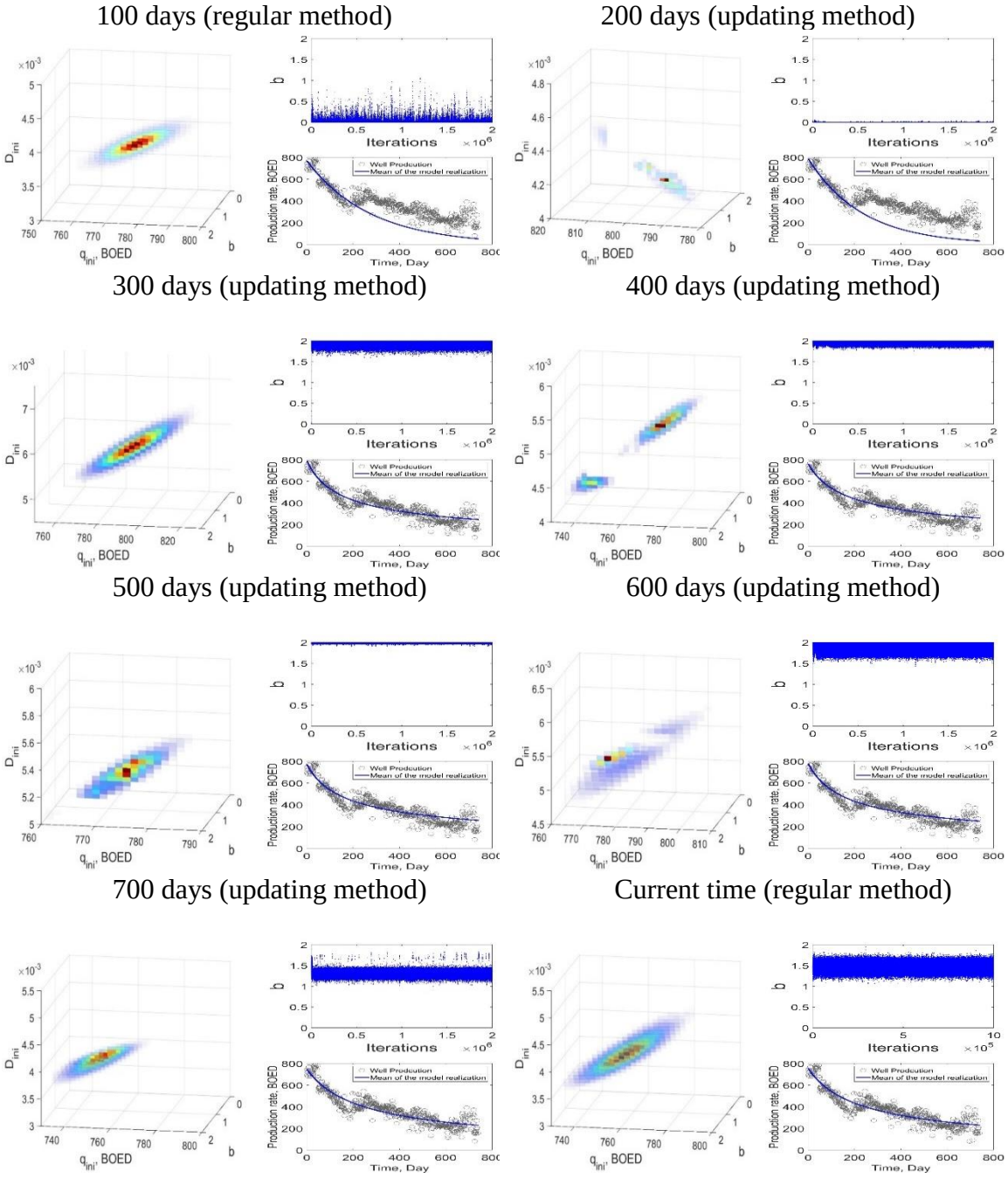

Fig. 6.5. The 3-dimensional posterior, the MCMC experiment of parameter ***b*** and the mean of realization of current time for updating and regular methods associated with well 1

The expected value of variable b for 700 days updating, varies from the 600 days and is similar to the current time regular analysis. However, the updating method permits an

uncertainty reduction in compare to the regular method for the same time interval. Also, an overview of the MCMC experiment in all time steps exhibits the inconsistency in the order of magnitude together with the heteroscedasticity of random parameter b. Moreover, it is worthwhile to note that, the comparison of posterior plots amid several updating time intervals, denotes the efficiency and robustness of the hybrid prior mechanism.

The gradual evolution of updating methods' posterior space associated with all times intervals are demonstrated in a same-scale formation (Fig. 6.6). The plot allows to comprehend the overall performance of updating data in the model results coupled with the state of certainty when more data becomes available. It is observed that, despite the correlation structure among variables pursue a similar trend the magnitude of it is altered case by case.

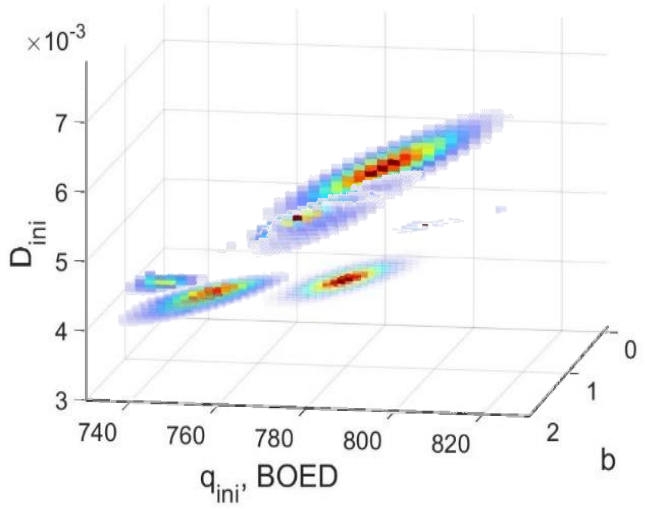

Fig. 6.6. Comparison plot of posteriors regarding the regular and updating methods

Another important feature which should be addressed in detail, is the comparison of the last updating step with the regular analysis of current production time. The MCMC

experiment of whole random field as well as the cumulative mean (C.mean) and cumulative standard deviation (C.std) is depicted in Fig. 6.7. The MCMC experiment of regular method (black dots) scattered more and is partially covered with the updating method (gray dots), which justifies the privilege of applying updating over regular Bayesian analysis. Additionally, results of cumulative mean and standard deviation plots, also supports utilizing the updating method, provided that the magnitude of variance appears smaller than the regular method in all experiments.

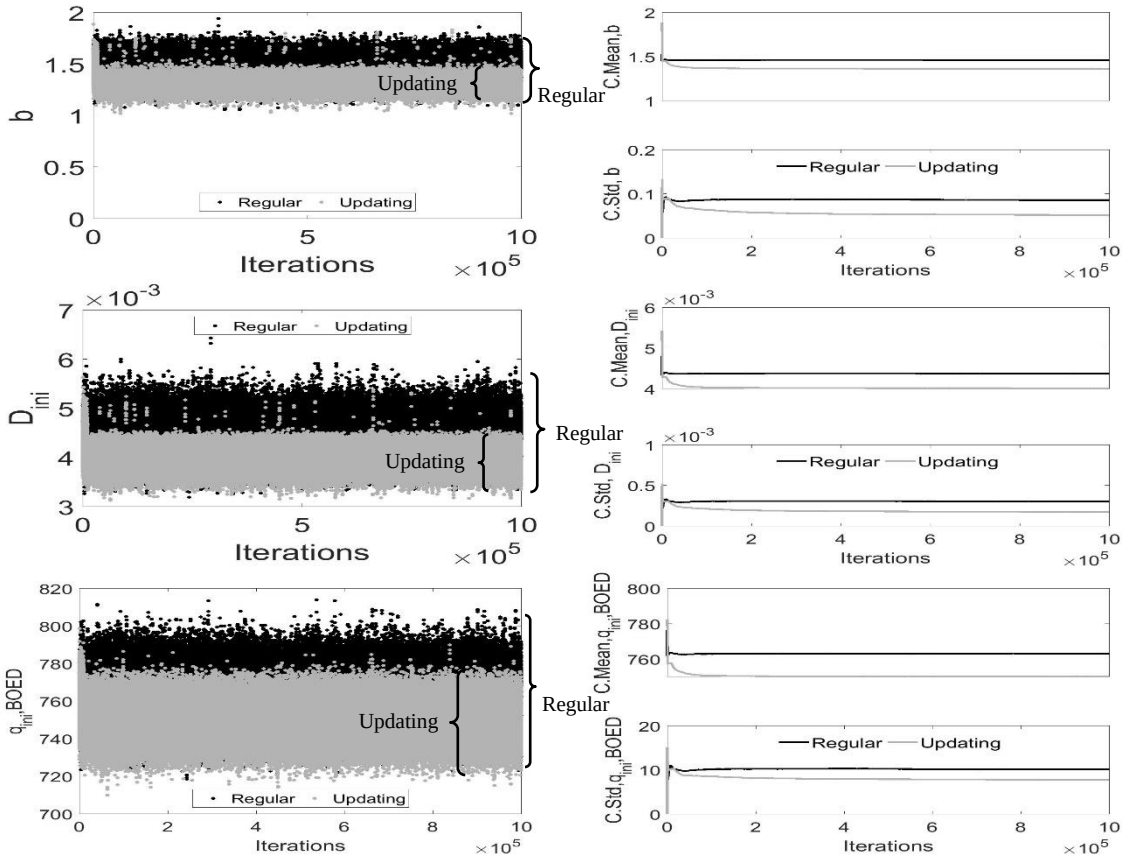

Fig. 6.7. Comparison plots of MCMC experiments corresponding with the last step updating and the current production time

The mean and standard deviation of model realizations for each individual time intervals of updating method along with the current production time of regular method are

coupled with the observed data and demonstrated in Fig. 6.8. The comparison of the mean of realizations implies the state of unreliability of model analysis when the well production is still at the early stages. Which is evident from the considerable deviation between the curvatures of 100 and 200 days with the other time intervals. Moreover, the comparison plot of standard deviation confirms the higher order of uncertainty in 100 days' results, while 500 days in contrary presents more confidence. Sequentially, the 700 days updating plot also provides the next higher computation certainty.

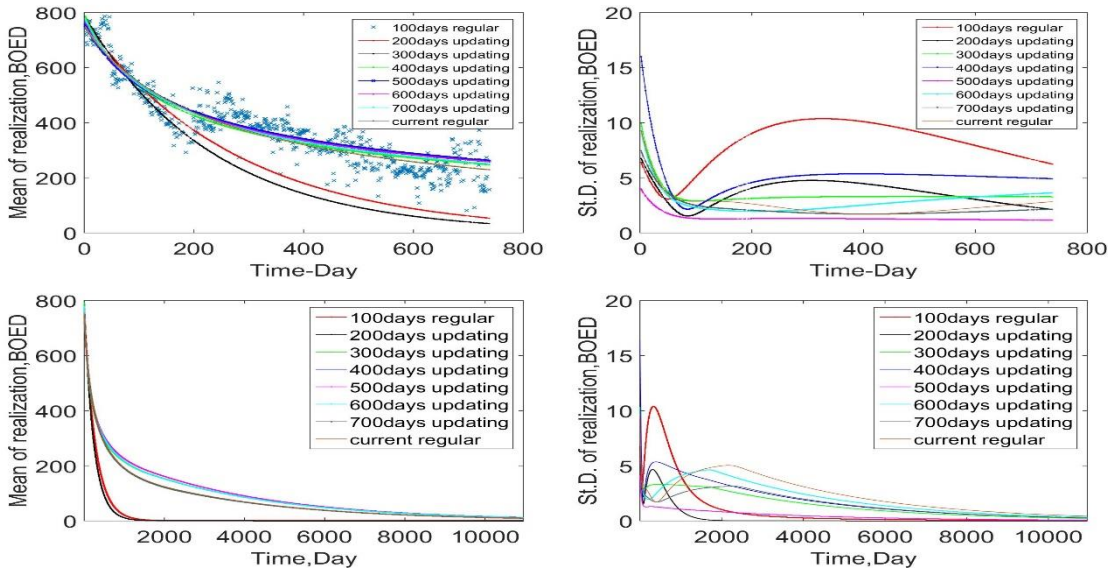

Fig. 6.8. Left, the mean and right, the standard deviation of realizations for updating time intervals together with the current production time and observed data. Top, current time and down, the extrapolation to 30 years depletion

It is a common practice to extrapolate the forward model outcome to construct the long-term production forecasting. Fig. 6.8, down, depicts the mean and standard deviation of 30 years' production associated with various time intervals. The model anticipation of 100 and 200 days, noticeably underestimate the well production. Provided that, the same result can be concluded from the standard deviation plot. Despite the level of confidence

become less in the 100 and 200 days results up to the 2000 days' prediction, by progressing in time, the current time production case induces more uncertainty into the statistical inference.

Fig. 6.9 outlines the significance of 700 days updating and regular Bayesian methods for whole experimental designs. The mean of realizations of both updating and regular Bayesian method, provide a subtle match with the observed data in all cases. However, the plot of standard deviation at right, is considerably informative and clearly defines the influence of applying different methods.

The standard deviation plot for Well 1, exhibits several conjunctions of regular and updating methods, while eventually, the updating method developed more certain analysis. Well 2 presents a distinguishable disagreement between two methods and the level of confidence in updating mechanism becomes substantially outweighed. The uncertainty in Well 3 is almost analogous, nevertheless, after 200 days of production, the updating method proceeds the regular framework. In agreement with other results, Well 4 demonstrates a similar behavior regarding the improvement in the order of confidence for updating method. In addition, the mean and standard deviation of realization of 30 years' production are depicted in Fig. 6.10. The general trend of mean of model realization in whole experimental cases indicates the homogeneity in together the methods of updating and regular. However, considering all experimental cases, the level of confidence undeniably is higher in the updating method.

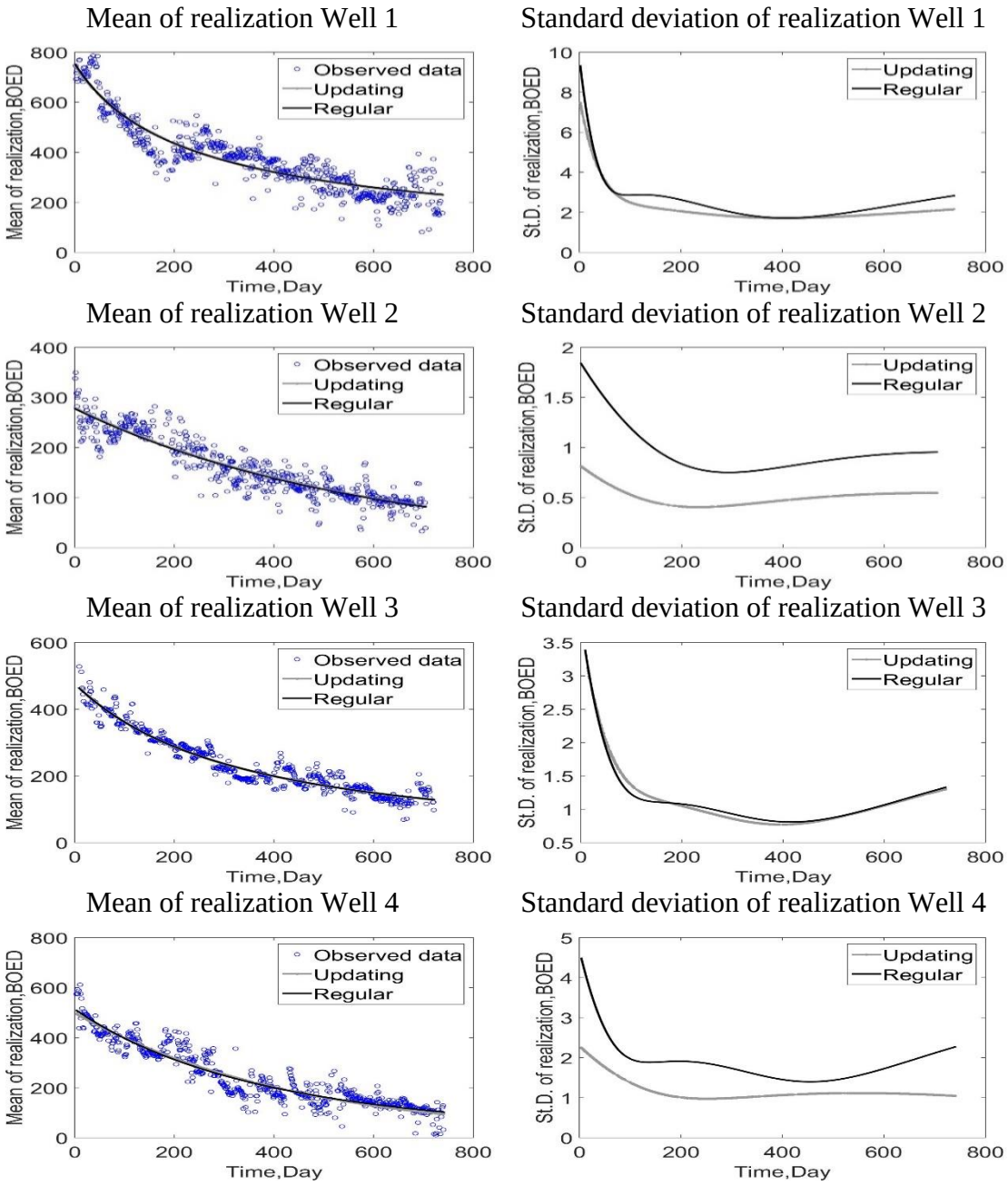

Fig. 6.9. The mean and standard deviation plots of together 700 days updating and regular methods corresponding to each case studies

In general, it is observed that, the updating method provide a constructive influence on total performance of the Bayesian analysis, when an appropriate prior is proposed.

The concise remarks of this section are visited in section 8.1.5.

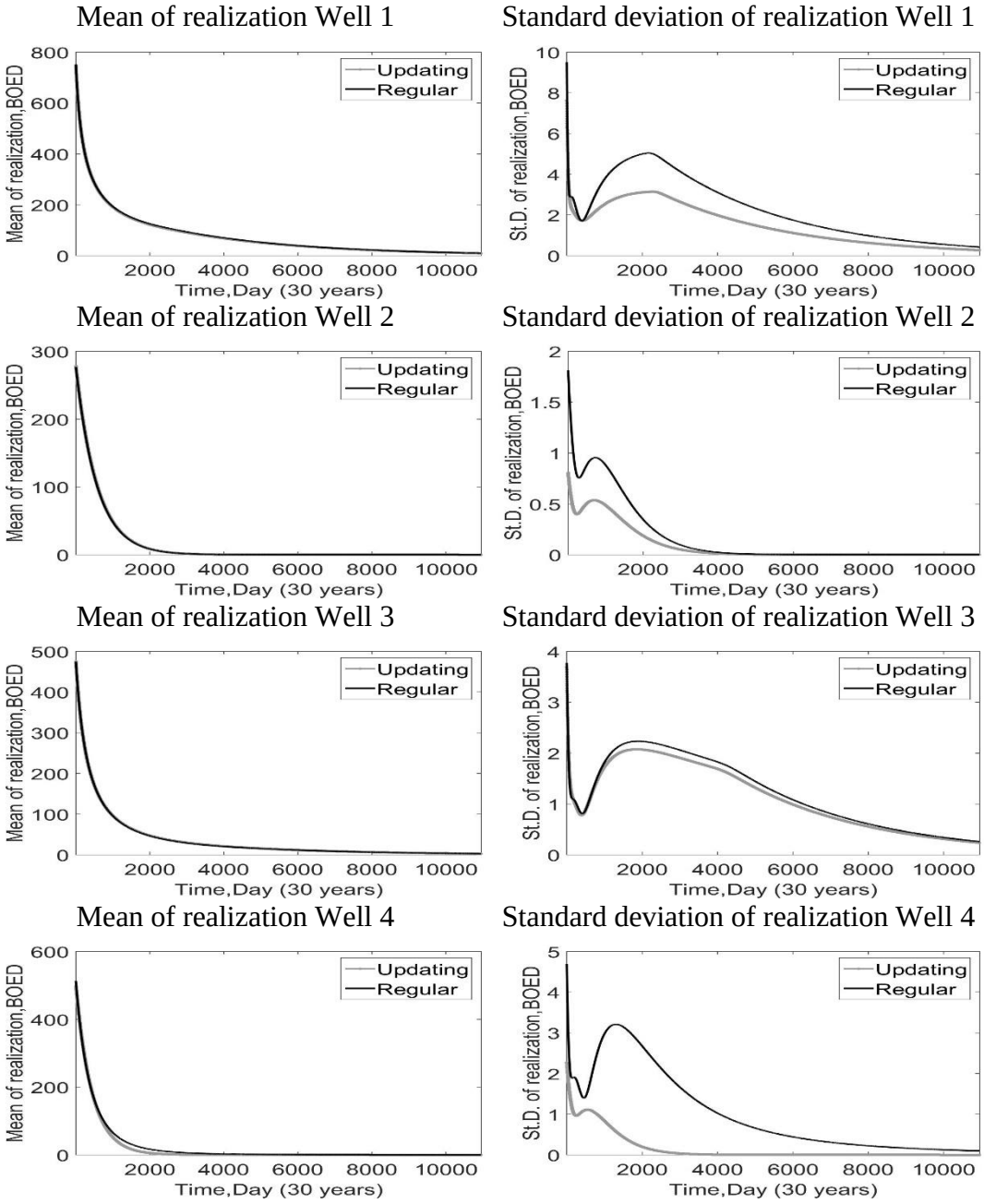

Fig. 6.10. The comparison plots of realizations mean and standard deviation regarding the 700 days' updating and regular frameworks for 30 years' production

7. SPATIAL DATA ANALYSIS OF UNCONVENTIONAL OIL AND GAS FORMATIONS: A BAYESIAN APPROACH

7.1. Overview

Applying spatial analysis frameworks allow us to approximate the short and long term Estimated Ultimate Recovery (EUR) for an unexplored location across a formation relying on the observed data at known locations. However, it has alluded that the EUR of wells at known locations are most often unidentified and should be retrieved exerting production forecasting models. Also, the likely associated correlation coupled with the unknown influence of incorporated distance amid the points makes it pertinent to employ a mechanism to quantify the uncertainty respected to the prediction model parameters as well as model realizations. Hence, that is for, the Bayesian paradigm is exploited as a mean to not only provide the inference of random field but also assess the uncertainty regarding the computational analysis.

Therefore, we elaborated a Bayesian-Spatial algorithm to constitute the spatial features of untouched locations hypothesizing the fact that the only given information encompasses the production observed data and corresponding coordinate for each individual well along with an appropriate EUR evaluation model.

In this study, the Power Law Exponential Decline (PLED) and Modified Hyperbolic Decline (MHD) curve methods serve to delineate the well production performance on the course of the progressing time. Additionally, the depletion logs of 43 wells, captured from

the Eagle Ford Shale demonstrate the observation data required to generate the Bayesian inference.

Moreover, the Markov Chain Monte Carlo (MCMC), Metropolis-Hastings (MH) and Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) as the subordinate algorithms of the Bayesian approach are used to sample the random field by imposing a suitable acceptance-rejection criterion. In addition, in order to deal with the spatial analysis, two techniques comprising the Ordinary Kriging (OK) method along with the Inverse Distance Weight (IDW) are used and compared together. To address the identified problem, initially applying the Bayesian probabilistic approach, the first and second order statistics of model parameters altogether with the vector of expected and variance of model realizations for the short and long term of production are retrieved. Next, these data coupled with the wells' coordinates feature the required information to establish the spatial analysis.

It is eventually implied that, given merely the observation data, associated coordinates and EUR evaluation models are enough to estimate the model variables and the production behavior for different courses of time at desired locations. Comparing different spatial analysis techniques, it has appeared that the OK-Exponential and then OK-Spherical models exhibited better forecasting results with substantially less associated standard errors.

7.2. Introduction

The spatial analysis methods have been applied as practical practices to predict various characteristics of interest field across the aim space (Ye Zhang 2009). The techniques employed to approximate the target quantity often allow to project the

influence of given estimates for available points considering the proportional impact of the associated distance amidst the current and desired locations (Isaaks and Srivastava 1989). Furthermore, the Ordinary Kriging (OK) and Inverse Distance Weight (IDW) are the local spatial analysis methods that due to their flexibility in implementation are selected as the reference spatial methods. Wackemagel, (2003) thoroughly discussed the available Kriging methods and their applications. Weber and Englund, (1992) and Lu and Wong, (2008), also separately featured various techniques to implement the Inverse Distance Weight (IDW). A comparison between the Ordinary Kriging and Inverse Distance Weight regarding the contributed chemical specifications of Soil is expanded in the Gotway et al., (1996) and Yasrebi et al., (2009). Moreover, Pyrcz and Deutsch, (2014) comprehensively elucidated several spatial techniques and their applications in the field of oil and gas developments. Additionally, a novel Kriging technique to spatially evaluate the permeability of rock mass in a reservoir is introduced in Brown and Falade, (2003). The correlation amid the geological characteristics and Kriging method is also discussed in Zhang et al., (2005) and Tian et al., (2017).

Markov Chain Monte Carlo (MCMC) coupled with the Metropolis-Hastings (MH) algorithm (Gelman et al. 2013)are subordinates of the Bayesian analysis that develop a reliable mechanism to sample the random field by utilizing a rejection-acceptance criterion. Nevertheless, it is evident that the MCMC sampling usually suffers from being computationally inefficient, which justifies applying more sophisticated methods such as the Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) framework (section 2.3.2). PSAMH is an augmented framework that ensures capturing all plausible posterior modes

together with the computational time reduction by exerting several synchronous chains and convergence enhancement features to optimize the acceptance rate. The retrieved posterior vector from PSAMH, permits computing the set of expected values and standard deviation of model realizations for the short and long terms of productions (section 4 and section 6). The standard deviation, hence, later can be used to quantify the uncertainty corresponding to the EUR analysis.

In addition, the Modified Hyperbolic Decline (MHD) and Power Law Exponential Decline (PLED) functions are two Modified-Arps' models that are used to evaluate the Estimated Ultimate Recovery (EUR) of each individual well with the unit of the Barrel of Oil Equivalent per day (BOE/D). The aforementioned frameworks which hereafter are evoked as "forward model" have lately become popular among the scholars and practitioners' due to their simplicity in the implementation together with the extended application in either conventional or hydraulic-fractured wells. We separately constituted the implementation procedures along with the application of MHD and PLED methods in the Eagle Ford Shale considering the regular and updating Bayesian probabilistic approaches in sections 4 and 5, and also Moridis et al., (2017). Also, Deutsch and Zanon, (2007), Willigers et al., (2014), Al-mudhafar et al., (2015) and Tarrahi et al., (2016) developed the notion of exploiting the Bayesian approach in the reservoir spatial analysis by providing various techniques when the data requires to be spatially evaluated is initially known.

In this study, by hypothesizing that the observed data, associated coordinates and forward models are the only available information, we establish a Bayesian mechanism to

attain the statistics of forward model properties and well productions' realizations for the progressing course of time. Furthermore, exerting the retrieved data, allows to derive the spatial analysis across the formation by inserting the first and second order statistics respected to forward model parameters as well as the daily basis and cumulative production for different time intervals, which we call it the dynamic mapping.

7.3. Methodology

Initially, we briefly introduce Modified Hyperbolic Decline (MHD) and Power Law Exponential Decline (PLED) models, which are discussed in section 4.3.1 and 4.3.2. Also, the definition of Markov Chain Monte Carlo (MCMC) and Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) are thoroughly elaborated in sections 2.3 and 2.3.2. Techniques of spatial analysis comprising the Inverse Distance Weight (IDW) and Ordinary Kriging (OK) is discussed in subsequent.

7.3.1. Spatial data analysis

We employed two well-known spatial analysis methods comprising the Inverse Distance Weight (IDW) and Ordinary Kriging (OK), which are shortly described in this section.

7.3.1.1. Inverse distance weight (IDW)

Inverse distance weight is a common practice to implement the spatial analysis, specifically in the case of point analysis (Weber and Englund 1992; Isaaks and Srivastava 1989; Ye Zhang 2009).

$$Z_0 = \frac{\sum_{ii=1}^{m} Z_{ii} \cdot \Delta_{ii}^{-\beta}}{\sum_{ii=1}^{m} \Delta_{ii}^{-\beta}} \tag{7.1}$$

Where, Z_0 and Z_{ii} ($ii \in [1, m]$) indicate the estimation value at new coordinate and the sample value at point ii, respectively. In addition, Δ_{ii} denotes the distance sample point to estimated point, while β presents the power factor to determine the weight influence (in this study, after several trial and errors β sets as 2.5).

7.3.1.2. Ordinary Kriging

In order to derive the Ordinary Kriging (OK), several steps should be implemented which are briefly described subsequently.

Experimental Variogram

Eq. (7.2) allows us to assess the state of correlation, $\omega(h)$, between the current location values (Isaaks and Srivastava 1989; Bohling 2005a) at constant distance intervals, h.

$$\omega(h) = \frac{1}{2m(h)} \sum_{ii=1}^{m(h)} [z(x_{ii} + h) - z(x_{ii})]^2 \tag{7.2}$$

Where, $m(h)$ indicates the number of sample pairs within the distance interval h. Also, $z(x_i + h), z(x_i)$ manifest the samples' values at two points separated by the distance interval h.

Semivariogram

Providing the experimental variogram, several semivariograms (Bachmaier and Backes 2008) models can be employed to project the trend of the experimental variogram. Exponential, Spherical, Gaussian and Stable models (Wackemagel 2003) are used to

approximate the component of variograms. The corresponding mathematical formulation of models are given in Eq.(7.3) to Eq.(7.6).

Spherical model

$$\omega(h) = \begin{cases} nugget + sill\times(1.5\frac{h}{range} - 0.5(\frac{h}{arange})^3) & h \le range \\ nugget + sill & otherwise \end{cases} \quad (7.3)$$

Exponential model

$$\omega(h) = nugegt + sill\times(1 - \exp\left(-\frac{3h}{arange}\right)) \quad (7.4)$$

Gaussian model

$$\omega(h) = nugget + sill\times(1 - \exp\left(-\frac{3h^2}{arange^2}\right)) \quad (7.5)$$

Stable model

$$\omega(h) = nugget + sill\times(\tilde{\beta}(1 - \exp(-3\left(\frac{h}{range}\right)^{\tilde{\beta}})) \ , \ 0 \le \tilde{\beta} \le 2 \quad (7.6)$$

Where, $sill$ and $nugget$ define the maximum variogram and the intercept of the semivariogram model with $\omega(h)$ axis, respectively. $range$ denotes the maximum effective distance from the target location

OK approximation at new coordinate

Let $z(x_{ii})$ represents the random function and x_{ii}, presents the sample locations. The prediction for the new coordinates with unknown properties is given in Eq. (7.7).

$$z(x_0) = \sum_{ii=1}^{m} \lambda_{ii} \, . \, z(x_{ii}) \quad (7.7)$$

Where, λ_{ii} is the assigned weight to each single observed data and must fulfill Eq. (7.8), which turns to reckon the Lagrange parameter.

$$\sum_{ii=1}^{m} \lambda_{ii} = 1 \tag{7.8}$$

And the weights, λ_{ii}, are retrieved from the subsequent matrix equation.

$$\lambda_{ii} = K^{-1}k \tag{7.9}$$

Where, K and k correspondingly indicate the square matrix of semivariograms between the known data locations and a vector of estimated semivariograms within the data locations and the new target coordinate. Eq.(7.10) and Eq.(7.11) exhibit the mathematical interpretation of K and k.

$$K = \begin{bmatrix} \omega(x_{ii,jj}) & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 0 \end{bmatrix} \qquad ii\ , jj\ \in [1, m] \tag{7.10}$$

And

$$k = \begin{bmatrix} \omega(x_{0,ii}) \\ \vdots \\ lagrange(x_0) \end{bmatrix} \qquad ii\ \in [1, m] \tag{7.11}$$

$\omega(x_{ii,jj})$ and $\omega(x_{0,ii})$ denote the semivariogram weights of previously known locations together and know with the target locations, respectively. Note that, the kriging weights and covariance are entirely determined by the configuration of locations and the shape of the covariance model (semivariogram) and not by the values associated with the known locations. The significance of provided locations' quantities only appears in the experimental variograms to determine the appropriate semivariogram model.

Ordinary Kriging not only permits to compute the estimated quantity at desired location, but also the respected variance. Accordingly, Eq. (7.7) expands to a practical form of Eq.(7.12) (Bohling 2005b).

$$z(x_0) = \sum_{ii=1}^{m} \lambda_{ii} \,.\, z(x_{ii}) = \sum_{ii=1}^{m} \lambda_{ii} \,.\, z(x_{ii}) + (1 - \sum_{ii=1}^{m} \lambda_{ii}) \times mean(x_{ii}) \tag{7.12}$$

It should be taken into consideration that $mean(x_{ii})$ estimates the mean of known locations less than the $range$. If the distance exceeds beyond the $range$ somehow an effective distance other than $range$ should be determined to be able to compute the mean of locations lay inside the effective distance, otherwise the mean substitutes by the mean of entire available locations. Furthermore, the variance of OK estimation becomes available through Eq.(7.13).

$$var(x_0) = sill - \sum_{ii=1}^{m} \lambda_{ii} \,.\, k - Lagrange(x_0) \tag{7.13}$$

7.3.2. Cross-validation

Cross-validation permits to assess the performance of a spatial analysis routine by providing a metric to approximate the error at proximity of each location. In order to implement the cross-validation technique, the location which the corresponding value is already known is removed and utilizing the desired spatial analysis framework the associated quantity is estimated and it should be iterated to the last location. Then, retrieving the residuals between the predicted and known quantities the standard error (Everitt and Skrondal 2010) of all known locations can be evaluated. The standard error (SE) can be computed exploiting equation (7.14).

$$SE = \frac{\sqrt{\sum \varepsilon^2}}{n} \tag{7.14}$$

7.4. Observed data

A set of 43 well depletion logs from the Eagle Ford Shale formation, obtained and assigned as the observed data. Fig. 7.1 demonstrates the spatially arrangement of wells for the real production time (current) of well together at top in both normal and log-log scales. It is mentioned earlier that values of n in PLED and b in MHD are substantially sensitive to the production time. Hence, it makes it pertinent to set an analogous production time for all wells. By examining the set of wells, it turned out that 350 days is the minimum production time for several wells, hence, 350 days' production time set as the fix time interval for all wells (Fig. 7.1, down). Also, considering the raw data, most often, some degree of heteroscedasticity is observed that makes it necessary to filter the data. We employed the systematic filtering mechanism constructed and expanded in section 0to remove the outliers from the raw data in each single well prior the MCMC experiment (Fig. 7.1, down).

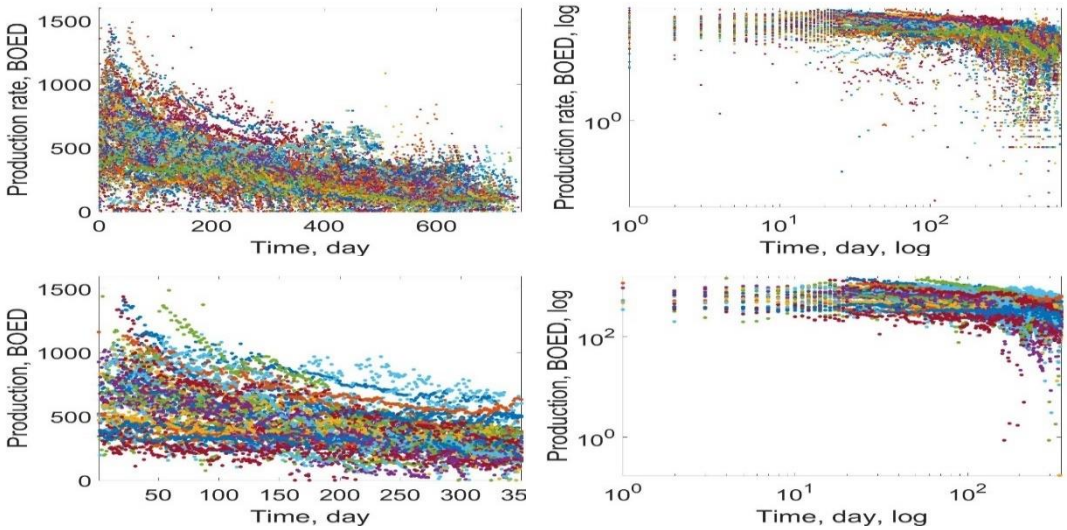

Fig. 7.1. The production rate of 43 wells from the Eagle Ford shale top, the real and down, 350 days production time in left, normal and right, log-log scales

7.5. Configuration of wells

Fig. 7.2 depicts the arrangement of wells ensemble in the relative distance measured in feet. According to the plot of well locations, it is observed that, despite some wells are in the proximity of each other, the others are extensively far and develop less influence on the other wells which suggests clustering of the wells ensemble.

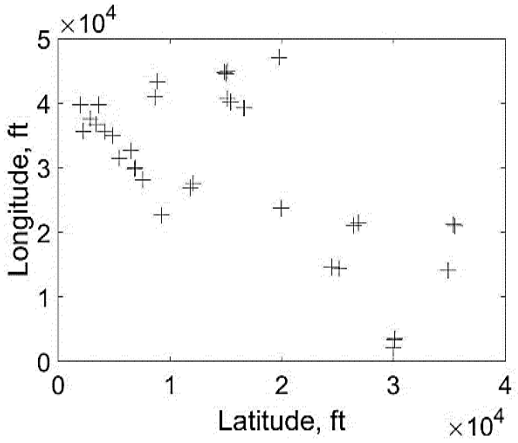

Fig. 7.2. The arrangement of 43 wells with the relative distance in feet

Moreover, we applied the Haversine formula (Van Brummelen 2012) to convert the distance between two locations to feet given the coordinates in Degrees.

In order to assess the distance, firstly it is essential to convert the degrees, minutes, seconds' coordinates to the decimal degrees.

$$Decimal\ Degrees = Degrees + \frac{Minutes}{60} + \frac{Seconds}{3600} \tag{7.15}$$

Additionally, exerting $= \frac{D\,\pi}{180}$, we would be able to convert the Degrees (D) to Radiant (R). Providing the coordinates in Radiant, now exploiting Haversine equation the distance (Δ) amid two locations can be approximated via Eq. (7.16) and Eq.(7.17).

$$\tau = Sin^2\left(\frac{\varphi_{2-latitude} - \varphi_{1-latitude}}{2}\right) + Cos(\varphi_{1-latitude})Cos(\varphi_{2-latitude})Sin^2\left(\frac{\varphi_{2-longitude} - \varphi_{1-longitude}}{2}\right) \quad (7.16)$$

Where, $\varphi_{,-,}$ indicates the latitude and longitude coordinates in Radiant corresponding to either target locations. Eq. (7.17) defines the distance within two coordinates in the desired unit.

$$\Delta = 2r\ arcsine\ (\sqrt{\tau}) \quad (7.17)$$

Where, r denotes the radius of Earth which respectively in Kilometer, Mile and Feet adopts the values of 6371, 3959 and 20,903,520.

7.6. Experimental design

In this study, there are two main objectives. The first task is to derive the vector of posterior together with the daily basis and cumulative realizations of models' production for short and long-term period of time. The next task is deliverable by applying the spatial analysis frameworks.

7.6.1. Task 1

To address the first task, it is required to generate the PSAMH sampler for each 43 well using two empirical models. The list of incorporated variables coupled with the experimental cases illustrated in Table 7.1.

Max (data) denotes the maximum value of production rate for each well. In addition, it can be alluded that there are altogether 86 $(43\times2 = 86)$ experimental cases retrieved from Task 1.

7.6.2. Task 2

To deal with featuring the spatial analysis study cases, considering the arrangement of well locations, we hypothesized two scenarios.

Table 7.1 Range of variables for experimental cases

Forward model	Variable	Range of variables [min, max]	Case Studies $ii \in [1,43]$
MHD	b	[1e-4, 2]	$C_{PSAMH-MHD-Well(ii)}$
	D_i	[1e-4, 20]	
	q_i (BOED)	[0.5, 2.5] ×max(data)	
PLED	n	[1e-4, 2]	$C_{PSAMH-PLED-Well(ii)}$
	D_i	[1e-4, 20]	
	D_∞	[1e-14, 2e3]	
	q_i (BOED)	[0.5, 2.5] ×max(data)	

An ensemble of all wells and an aggregate of clusters encompasses of two separated units that allow us to investigate the state of errors regarding the effective distance between wells. Fig. 7.3 demonstrates all wells ensemble as scenario one at left and the assimilation of two assumed clusters, Cluster 1 and Cluster 2 comprising 28 and 15 wells, respectively as scenario two at right.

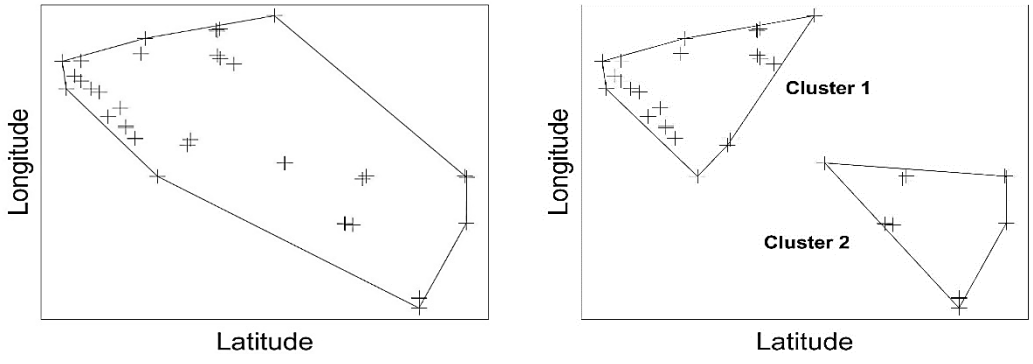

Fig. 7.3. Left, scenario one, all wells together and right scenario two, assimilation of two clusters

Furthermore, for each individual scenario the experimental studies contain several subcases as follows:

- The empirical models comprise 3 and 4 parameters corresponding to MHD and PLED, respectively.
- Dynamic production configuration of daily and cumulative productions for 1, 5, 10, 15, 20, 25 and 30 years of production.

To elucidate the total number of Task 2 experimental case studies, associated with each above aforementioned case, a special case study's name is introduced to be able to thoroughly delineate the intention of the experiment design (Eq. (7.18) and Table 7.2).

$$C_{(1)-(2)-(3)-(4)-(5)} \tag{7.18}$$

Table 7.2 illustrates the definition along with the initiation of subscripts indicated in the case study name.

Table 7.2 List of definitions of subscripts in the experimental design

Empirical model	Bayesian inference	Spatial Method	Statistics	Scenario
Modified Hyperbolic (MHD)	Decline curve model parameters (MP)	Inverse Distance Weigh (IDW)	Mean (M)	All
Power Law Exponential ($PLED$)	Production, daily basis (PD)	Ordinary Kriging-Spherical (OK_{Sp})	Standard Deviation (StD)	Cluster 1 ($C1$)
	Production, cumulative (PC)	Ordinary Kriging-Exponential (OK_{Ex})		Cluster 2 ($C2$)
		Ordinary Kriging-Gaussian (OK_{Ga})		
		Ordinary Kriging-Stable (OK_{St})		

For instance, $C_{MHD-MP(q_i)-IDW-M-All}$ initiates the case study regarding the mean (M) of Modified Hyperbolic Decline (MHD) model parameter (q_i) applying the inverse distance weight (IDW) as the spatial analysis method considering all wells ensemble scenario (All). Another example, $C_{PLED-PC(15)-OK_{St}-StD-C1}$ indicates the case study for the standard deviation (StD) of 15 years' cumulative production (PC (15)) of Power Law Exponential Decline ($PLED$) model exerting the Ordinary Kriging-Stable (OK_{st}) spatial analysis method relevant to cluster 1($C1$).

Fig. 7.4 demonstrates the diagram of different incorporated terms to be able to better perceive the components of case studies.

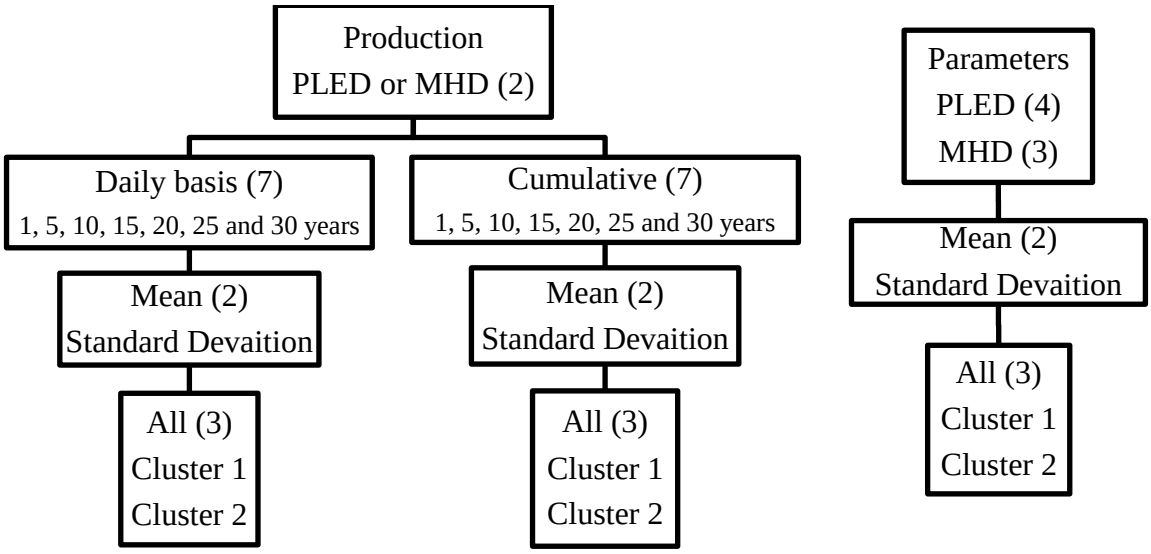

Fig. 7.4. The diagram of experimental design for dynamic production mapping (daily and cumulative) together with the decline curve models' parameters

The values inside parentheses (.), identifies the number of possible permutation in each case. Note that, regarding the definition provided in Table 7.2 and the diagrams in

Fig. 7.4, in total, 210 $(2\times[(7\times2\times3)+(7\times2\times3)]+3\times2\times3+4\times2\times3)=210)$ experimental cases are constructed applicable to Task 2.

Also, bear in mind that obtaining the standard error of cluster 1 and 2, the aggregate of errors regarding assimilation of clusters is identified by the subsequent equation

$$SE_{Clusters}=\frac{\sqrt{(n_{Cluster\ 1}\ SE_{Cluster\ 1})^2+(n_{Cluster\ 2}\ SE_{Cluster\ 2})^2}}{n_{Total}} \tag{7.19}$$

Where, $n_{Total}=n_{Cluster\ 1}+n_{Cluster\ 2}$.

7.7. Results and discussion

The results of each individual task are delivered and briefly discussed through the following subsections.

7.7.1. Bayesian analysis and MCMC (Task 1)

We delineate the implementation sequence for MCMC experiment through merely case Well 20 MHD, because the rest of Task 1 cases resemble to the provided case.

7.7.1.1. Raw data filtering and likelihood diagnostic

Often there are outliers in observed data that unnecessarily disrupt the general trend of production projection. We introduced a systematic mechanism that automatically filters the raw data by diminishing the skewness and removing outliers in section 0. We avoid repeating the procedure here and encourage the reader to study the mechanism from the referred paper. Fig. 7.5 depicts a comparison amide the raw observed (red square) coupled with the filtered data (gray star) and the fit NLS curvatures in the regular and log-log scale (top right and left).

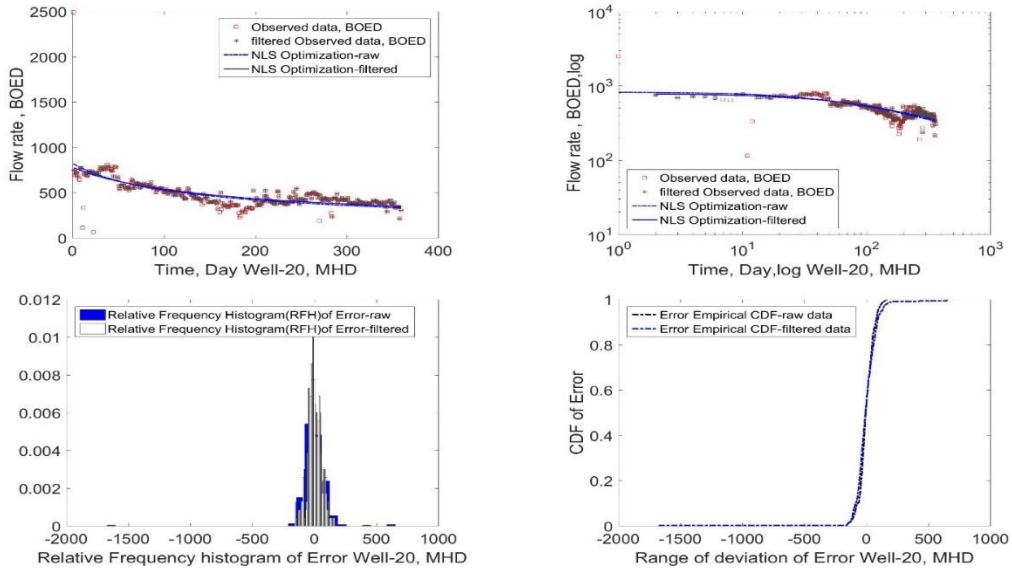

Fig. 7.5. Observed data before and after filtering top, left in the regular and right log-log scale. Relative frequency histogram and cumulative density function of raw and filter data down left and right, respectively

The comparison between the relative frequency histograms, down left, and cumulative density functions, down right, before and after filtering manifests the reduction of heavy left skewness exerting the filtering mechanism.

Moreover, the Nonlinear Least Square (NLS) optimization, which appeared in the filtering approach has become a key factor to recognize the plausible shape of likelihood from the residuals of observed data and optimized curvature. NLS, also, allows obtaining the initial values of random variables useful in MCMC. The best fit of Normal density function (red solid line) in both relative frequency histogram and the cumulative density function displays the precise selection of the Normal distribution as the likelihood (Fig. 7.6).

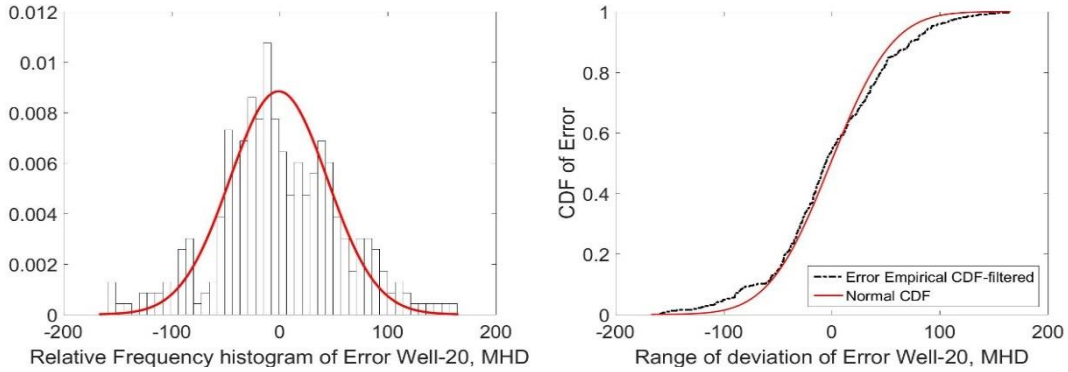

Fig. 7.6. Normal distribution fit to the relative frequency histogram of residual, left, and right the fit of cumulative density function of normal distribution to the residual

7.7.1.2. MCMC experiment

The retrieved optimized quantities associated with forward models' parameters permit to remove the ambiguity of initial values inserted into the MCMC and PSAMH sampler. The MCMC experiments for MHD model parameters are demonstrated in Fig. 7.7. In addition, the cumulative mean and standard deviation of random samples on the right side of Fig. 7.7 provides a tool to estimate the state of convergence of MCMC to assess the stationary of generated random samples. The MCMC is presumed to be converged when the cumulative mean and standard deviation of random field become straight lines after a specific iteration (Burn-In pint).

7.7.1.3. Statistics of posterior

While the first and second order of statistics can be attained from Fig. 7.7, right; Fig. 7.8 besides demonstrates the relative frequency histogram as well as the joint distribution of random variables. Although, the plot of relative frequency histogram of variable b depicts the tendency of the corresponding parameter to exceed beyond the boundaries, histogram of q_i and $D_i$clearly approximate the Normal distribution.

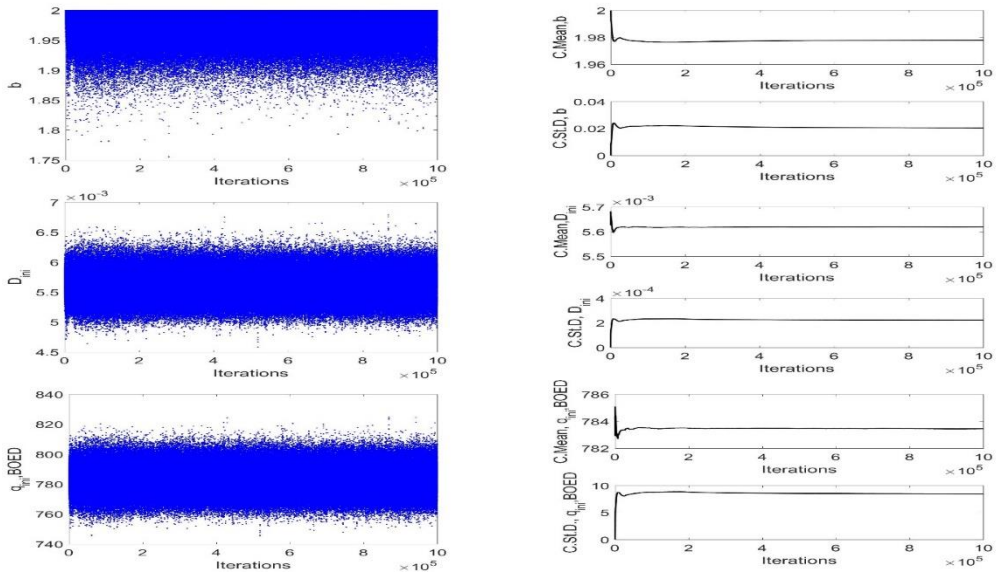

Fig. 7.7. MCMC experiment of MHD, left and cumulative mean and standard deviation, right

The joint distribution of q_i and D_i reveals the strong state of positive correlation, however, two other joint distributions present a partially non-correlated structure.

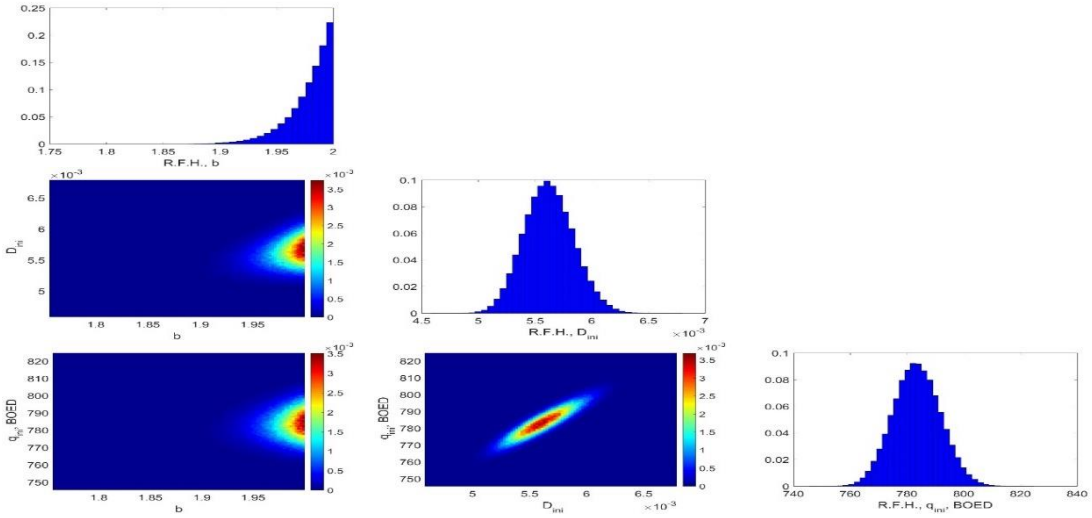

Fig. 7.8. Joint distribution with a side color bar and relative frequency histogram of forward model parameters

7.7.1.4. Current time realizations

Obtaining the vector of posterior allows us to draw thousands of realizations by inserting the pair of random samples into the forward models. Fig. 7.9 shows 10,000 realizations of daily basis and cumulative current production time.

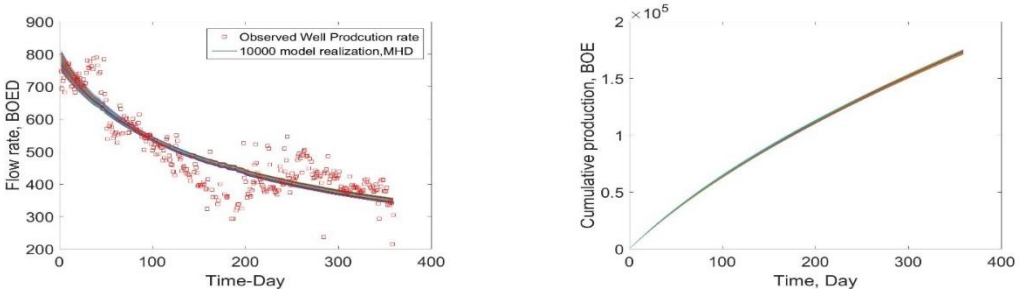

Fig. 7.9. 10,000 realizations of MHD fit to the observed data, left and associated cumulative realizations on right with respect to the 350 days

Fig. 7.10 depicts the mean of daily basis together with the cumulative production at top and the corresponding standard deviations at down. The standard deviation often interprets as the level of confidence regarding the computation and utilizes to quantify the uncertainty.

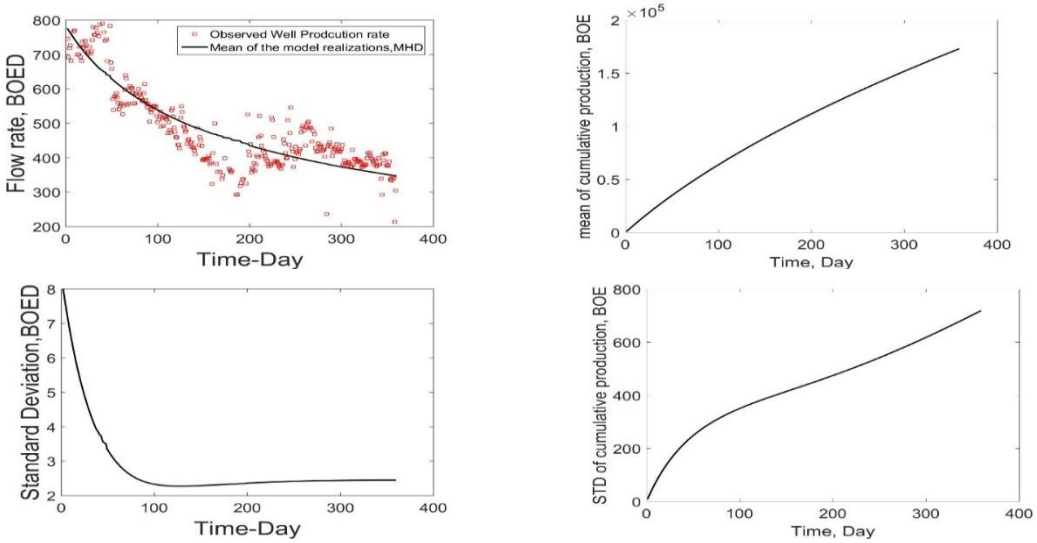

Fig. 7.10. The mean and standard deviation of realizations of both daily and cumulative productions for 350 days

According to the standard deviation plot of cumulative production, by progressing in time the level of confidence decreases.

7.7.1.5. 30 years' realizations

An extrapolation of inserted random set of posteriors incorporated in the forward model features the long-term production forecasting. Fig. 7.11 similar to the current time realizations, demonstrates 10,000 realizations in accordance with the 30 years extrapolation for daily and cumulative productions.

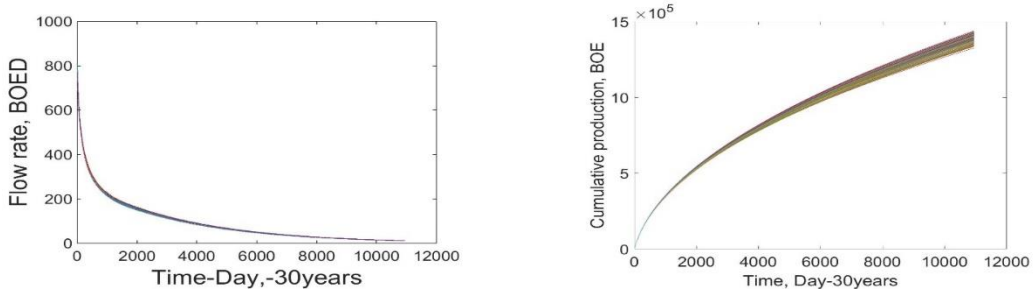

Fig. 7.11. 10,000 realizations of 30 years production in the daily and cumulative basis

The mean and standard deviation of 10,000 realization regarding the 30 years of production are presented in Fig. 7.12. The rate of growth of uncertainty drops after almost 2000 days of production (Fig. 7.12, down, left).

The mean and standard deviation of random fields altogether with the daily and cumulative productions associated with MHD and PLED are retrieved and then introduced as the input data of the spatial analysis which is initially defined as Task 2. Note that, the associated data are delivered in Appendix A.

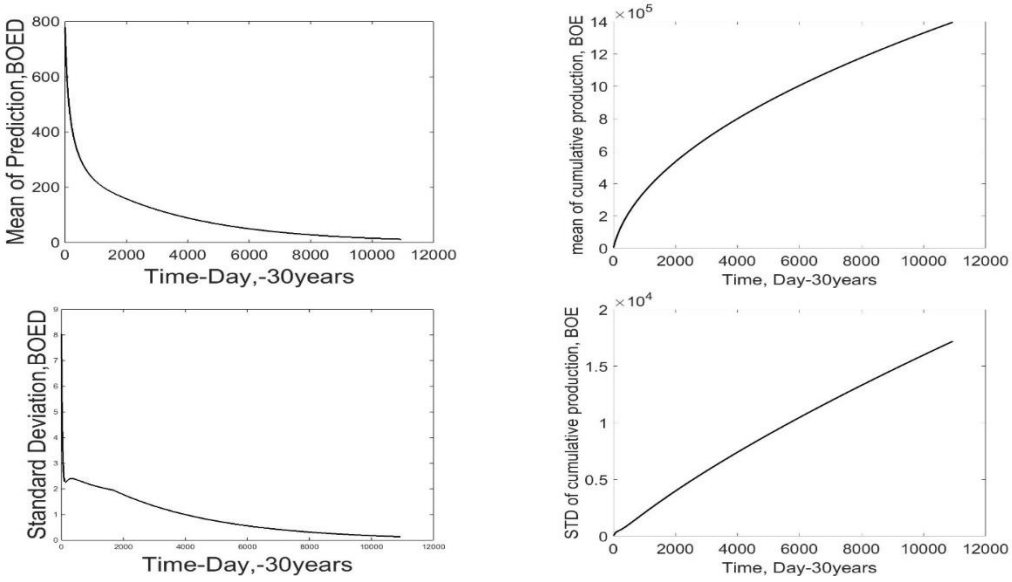

Fig. 7.12. Expected values and standard deviations of 30 years extrapolation in the daily and cumulative extraction

7.7.2. Spatial analysis (Task 2)

The spatial analysis comprises two major techniques, the Inverse Distance Weight (IDW) and Ordinary Kriging (OK) which is elaborated via the case of $C_{MHD-MP-,-M-All}$. The rest of computations is analogous to the provided case and pursue Task 2 experimental cases (Table 7.2 and Fig. 7.4). Furthermore, the configuration of MHD model parameters values retrieved from Task 1 are demonstrated in Fig. 7.13.

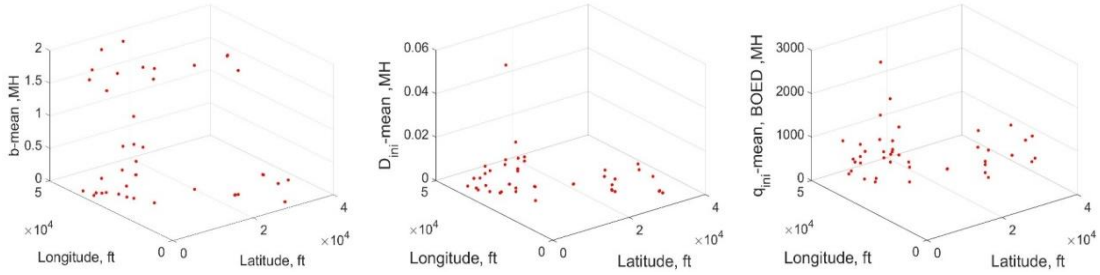

Fig. 7.13. Configuration of b, D_i and q_i quantities from left to right, respectively

7.7.3. Inverse Distance Weight (IDW)

Fig. 7.14 depicts the contour plots of b, $D_i$and q_i values derived from the IDW method. The contour of $D_i$implies that for a wide distance, IDW predicts similar range of quantities, however, contours of other variables convey more diversities.

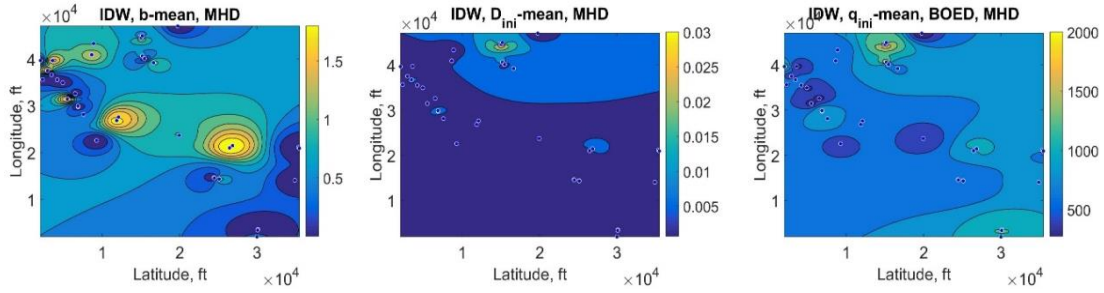

Fig. 7.14. Contours of IDW spatial method associated with MHD model parameters

7.7.4. Ordinary Kriging (OK)

It is alluded earlier that, four semivariograms are used in this study comprising the Spherical, Exponential, Gaussian and Stable models. Fig. 7.15 displays the experimental variogram and respected semivariograms. Thus, the semivariograms allow to incorporate the reckoned values of $Sill$, $Nugget$ and $Range$ into OK to assess the target location quantity.

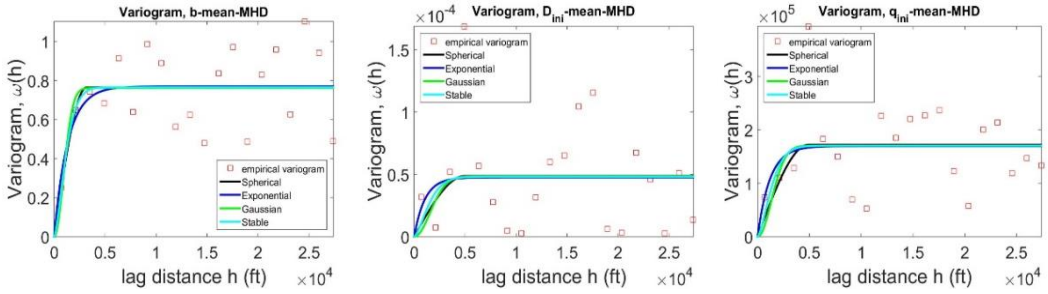

Fig. 7.15. Experimental and semivariogram associated with MHD mean model variables

The OK-Spherical analysis contours coupled with their variance for expected values of MHD model parameters are illustrated in Fig. 7.16. Note that, when the distance amidst two locations becomes larger than $range$ value, model simply computes the mean of neighbor locations and substitute it with the OK approximation.

That is for, intuitively it is evident that beyond a specific distance the mutual influence of locations becomes negligible and the rational option is replacing the estimation with the mean of proximate locations. The variance contours conduct the state of spatial correlation of OK estimation.

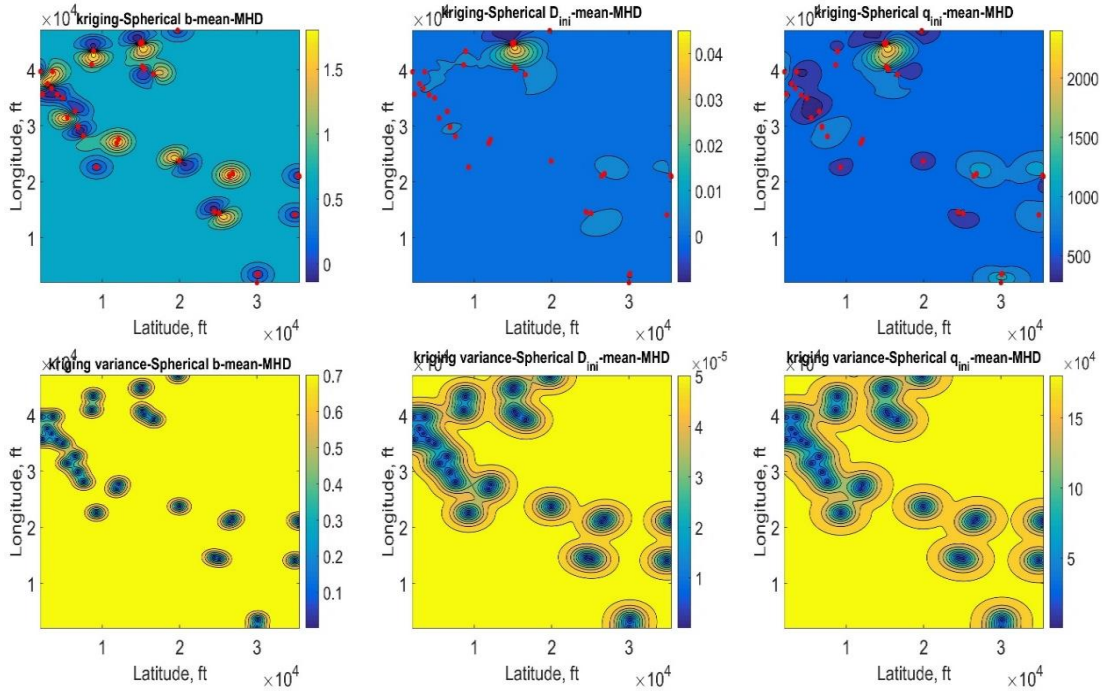

Fig. 7.16. OK-Spherical estimated value, top and variance contours for mean of MHD

According the presented plots, it is connoted that only locations close enough together provide a considerable impact which eventually justifies the notion of clustering.

The OK approximation continued with the OK-Exponential semivariogram model (Fig. 7.17). A comparison between OK-Spherical and OK-Exponential reveals the similarity between models. Observing the variance contours, despite OK-Exponential exhibits more dependency in parameter *b*, the correlation amid locations of two other variables almost disappeared.

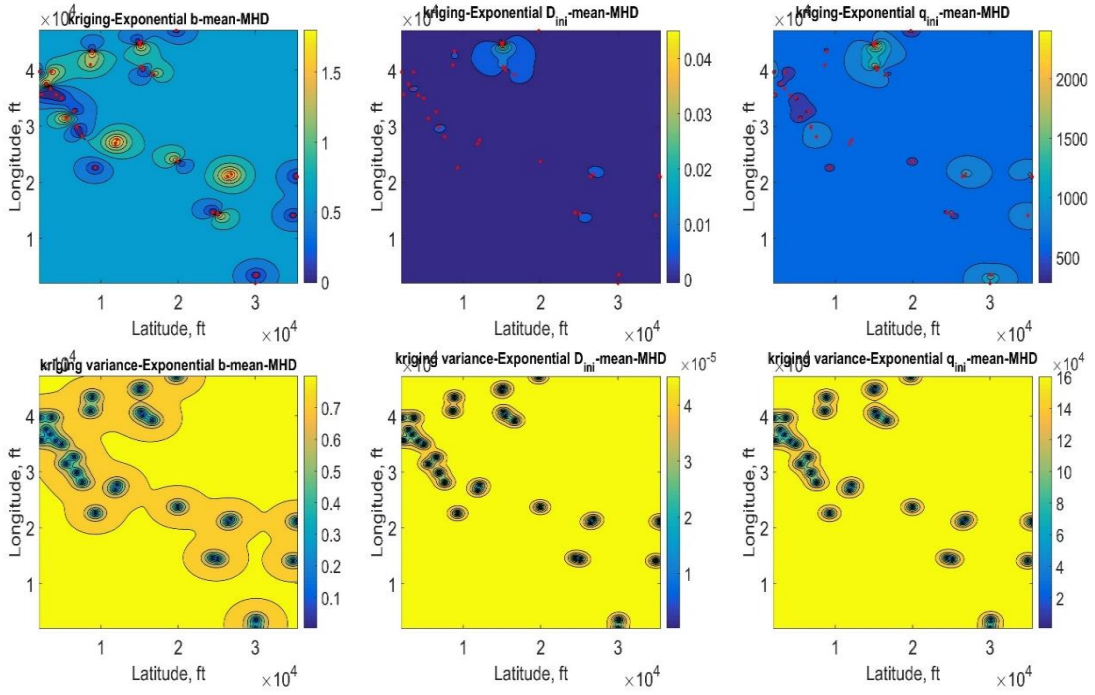

Fig. 7.17. OK-Exponential, estimated values, top, and associated variances, down

The OK-Gaussian semivariogram indicates a poor performance due to lack of enough correlation between locations depicted in Fig. 7.18. That is for, this model predicts negative quantities for MHD model parameters which is unrealistic.

Nevertheless, despite the poor performance regarding the target estimation, the variance of locations is in analogy with other visited models.

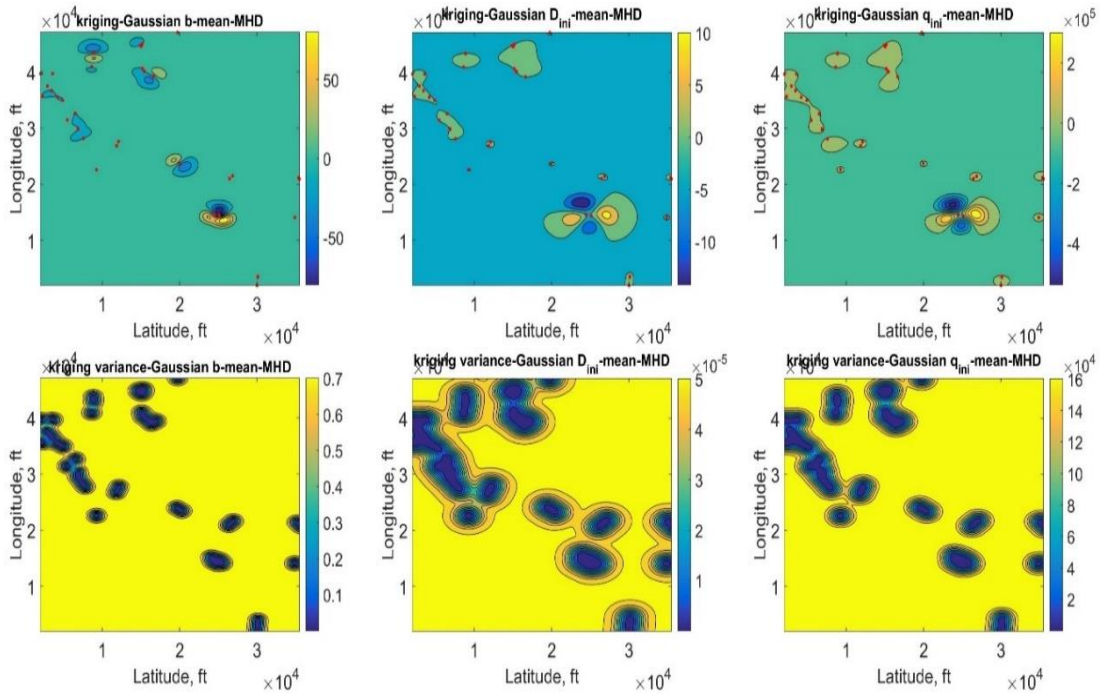

Fig. 7.18. OK-Gaussian semivariogram model for target estimations, top and down corresponding variances

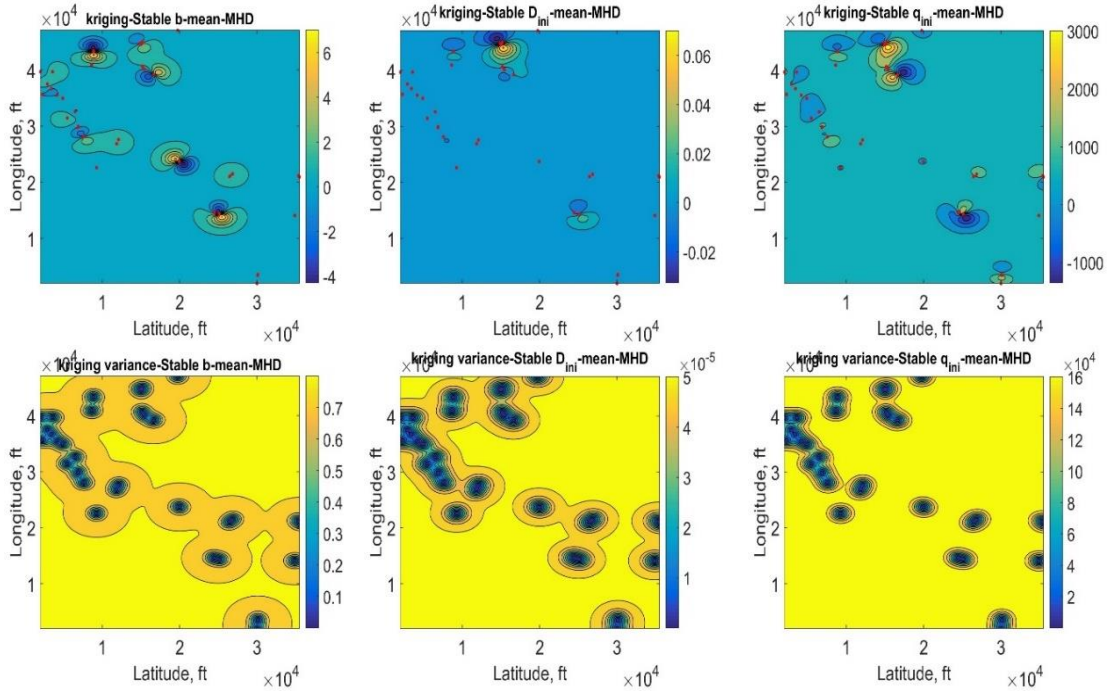

Fig. 7.19. Target location approximated by OK-Stable model top and respected variance, down

Acquisition of OK estimation leads us to the last semivariogram model which is OK-Stable (Fig. 7.19). OK-Stable is defined as a modification on the OK-Gaussian model by enabling manipulating the power of model. Considering the attained contours, it is implied that, the general trend of OK-Stable imitates the OK-Gaussian in the term of producing negative quantities. However, the order of magnitude of positive and negative values to some extent are smaller than OK-Gaussian.

7.7.5. Cross-validation results

Cross-validation appears as a reliable mechanism to evaluate the performance of spatial analyzing techniques. The computed standard errors (SE) associated with cross validation for case MHD parameter b are illustrated in Table 7.3 and rest of results in the form of two categories attributed with the empirical frameworks, MHD and PLED are presented in the Appendix B.

Table 7.3 Expected values coupled with the standard deviation of MHD parameter b

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.61E-01	1.20E-02	1.76E-01	1.35E-02	2.89E-01	2.48E-02	3.39E-02	2.79E-03
$\boldsymbol{OK_{Sp}}$	1.58E-01	1.08E-02	1.41E-01	1.24E-02	2.56E-01	2.48E-02	2.89E-02	2.71E-03
$\boldsymbol{OK_{ex}}$	1.56E-01	1.08E-02	1.42E-01	1.22E-02	2.63E-01	2.48E-02	2.94E-02	2.69E-03
$\boldsymbol{OK_{Ga}}$	3.36E-01	1.39E-02	1.42E-01	1.51E-02	5.30E+00	9.81E-02	4.78E-01	9.02E-03
$\boldsymbol{OK_{St}}$	1.58E-01	1.08E-02	1.40E-01	1.24E-02	2.78E-01	2.45E-02	3.04E-02	2.69E-03

Obtaining all case studies associated with Task 2 it becomes possible to draw several conclusions regarding the efficiency of spatial analysis techniques or the influence of empirical models' selection on the analysis performance. Subsequent, we address two inquiries; firstly, among two scenarios, Cluster and All coordinates together, which one

provides a more precise approximation. Secondly, to investigate which one of spatial techniques suggests minimum standard error (ES) in general.

7.7.5.1. MHD

Exerting the cross-validation method and using the Appendix B, allow comparing the impact of various spatial techniques on the course of model parameters.

Clusters or All coordinates

The minimum SE charts respected with the MHD model parameters utilizing IDW and OK models altogether are depicted in Fig. 7.20. The plots admit the advantage of clustering over all coordinates scenario, which in this case the clustering scenario accidently indicates 93.3% regarding both expected value and standard deviation of MHD model parameters.

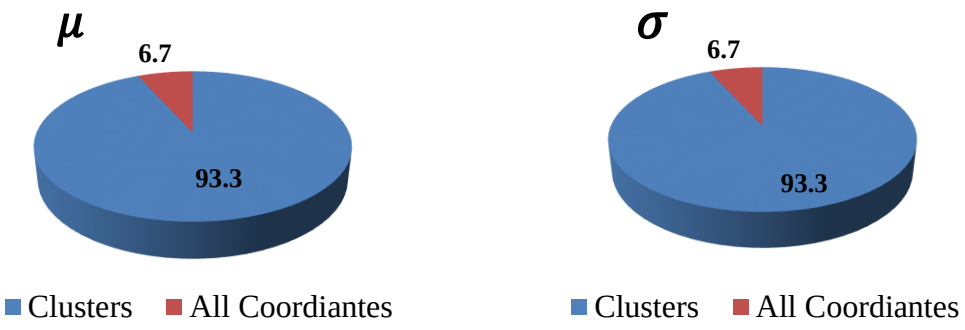

Fig. 7.20. Minimum SE of MHD model parameters, left expected value and right, standard deviation applying all spatial analysis methods

The impact of clustering in comparison to the all coordinate scenarios for MHD case studies are outlined in Table 7.4. In order to deal with the production either daily basis or cumulatively, the minimum standard error of entire time intervals integrated and illustrated in Table 7.4.

Table 7.4 Efficiency of clustering versus All coordinates in different scenarios

Experimental case	Model parameters		Daily basis Production		Cumulative Production	
Scenario	μ	σ	μ	σ	μ	σ
Clusters	93.3	93.3	100	98.6	100	97.1
All Coordinates	6.7	6.7	0	1.4	0	2.9

The provided results admit the substantial impact of clustering on the performance enhancement in MHD model by decreasing the standard error over both model parameters and productions.

Spatial analysis technique

Fig. 7.21 depicts comparison charts amid different spatial analysis techniques regarding the aggregation of MHD model parameters expected values and standard deviations. The results suggest that the OK-Exponential develops a better approximation of MHD model variables.

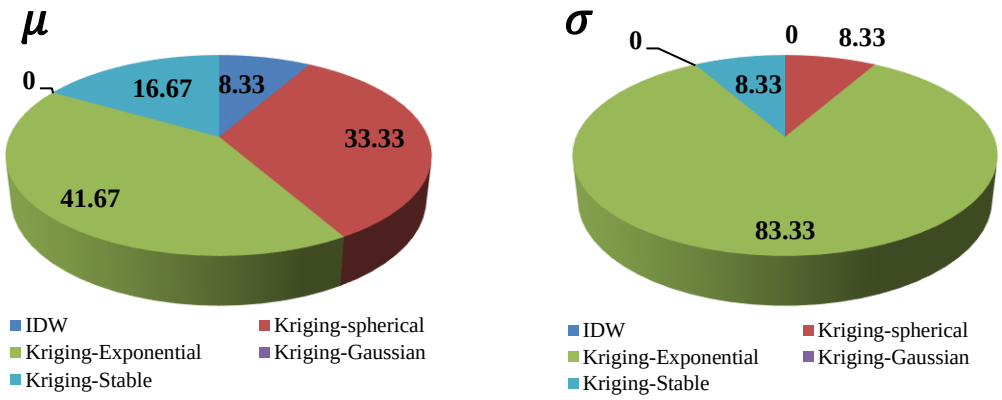

Fig. 7.21. The minimum SE of spatial analysis methods considering all MHD model parameters

In accordance to above charts, the contributing percentage of each technique in producing minimum standard errors is computed and illustrated in Table 7.5.

Table 7.5 Evaluation of spatial analysis techniques performance associated with the MHD model

Experimental case	Model parameters		Daily basis Production		Cumulative Production	
Technique	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	8.33	0	0	3.57	3.57	7.14
$\boldsymbol{OK_{Sp}}$	33.33	8.33	0	0	0	10.71
$\boldsymbol{OK_{ex}}$	41.67	83.33	82.14	96.43	92.86	82.14
$\boldsymbol{OK_{Ga}}$	0	0	14.29	0	0	0
$\boldsymbol{OK_{St}}$	16.67	8.33	3.57	0	3.57	0

The obtained results, undeniably indicates the advantage of OK-Exponential method versus other spatial analysis techniques.

7.7.5.2. PLED

The same conclusions can also be drawn for PLED model parameters.

Clusters or All coordinates

The efficiency percentage presented in Table 7.6, again proves the positive influence of clustering against all coordinates associated with PLED model parameters, daily and cumulative productions.

Table 7.6 Evaluation of effectiveness of clustering and all coordinates corresponding with PLED

Experimental case	Model parameters		Daily basis Production		Cumulative Production	
Scenario	μ	σ	μ	σ	μ	σ
Clusters	100	95	85.7	91.4	85.7	97.1
All Coordinates	0	5	14.3	8.6	14.3	2.9

Spatial analysis technique

Despite the fact that in some cases the efficiency of OK-Spherical model becomes close to the OK-Exponential in Table 7.7, the definite advantage belongs to OK-Exponential model.

Table 7.7 Impact of spatial analysis models regarding PLED model parameters and productions

Experimental case	Model parameters		Daily basis Production		Cumulative Production	
Technique	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	12.5	6.25	0	25	3.57	28.57
$\boldsymbol{OK_{Sp}}$	6.25	0	28.57	28.57	21.43	3.57
$\boldsymbol{OK_{ex}}$	81.25	81.25	67.86	46.43	71.43	67.86
$\boldsymbol{OK_{Ga}}$	0	12.5	0	0	0	0
$\boldsymbol{OK_{St}}$	0	0	3.57	0	3.57	0

7.7.5.3. Comparison of PLED and MHD

Another conclusion derived from Task 1 and Task 2 together is the assessment of performance of PLED versus MHD productions in the course of progressing time, daily basis or cumulatively, considering the produced standard errors (SE). In order to compute the respected SE, the minimum error retrieved from entire time intervals added together. The associated results in Table 7.8 indicates the irrefutable privilege of MHD method versus PLED in both experimental cases by generating less SE.

Table 7.8 Comparison between the MHD and PLED daily and cumulative production

Experimental cases	Daily basis production		Cumulative production	
Forward model	μ	σ	μ	σ
MHD	1.19E+01	6.34E-01	7.71E+04	5.20E+03
PLED	1.51E+01	1.55E+00	8.18E+04	8.42E+03

This section's conclusion is shortly recapitulated in section 8.1.6.

8. CONCLUSIONS AND FUTURE RESEARCH

The significant vertices of each section are separately and concisely recapitulated here. In addition, several alternative future researches are proposed proceed by conclusions.

8.1. Conclusions

8.1.1. Section 2

- Parallel Scaled Adaptive Metropolis-Hastings (PSAMH) is a practical method that aims to liberally explore the posterior space by incorporating the adaptive MCMC technique in a more efficient approach.
- In addition to the core algorithm, several features provided mechanisms to readily augment PSAMH by automatically tuning the step size of proposal distribution and removing the concurrent chains redundancy to achieve the optimum acceptance rate.
- Additionally, a synthetic case study is utilized to delineate the implementation sequence of PSAMH coupled with its application in the Bayesian inference via model realizations.
- Regardless of the type MCMC sampling method, considering either relative frequency histograms, cumulative density function plots, or model realizations' inferences, it is evident that by adding more observations the precision of analysis significantly improved (P10R1 to P10R5).

- Nevertheless, after a certain point, the rate of improvement in the model computation and uncertainty reduction are diminished, when means and standard deviations of model realizations for P10R5 and P10R10 produced approximately similar results.
- Moreover, plots of acceptance rate together with the scale factor provided extra tools to assess the state of stationary in addition to the online mean and standard deviation of random fields.

8.1.2. Section 23

- The use of a synthetic case study aimed to provide a mean to evaluate the performance of different well-known samplers in compare to PSAMH framework when the standard deviation and correlation coefficient values are varied.
- It is generally implied that PSAMH method not only accurately explores the random filed but also reduces the computation running time, and hence, increases the efficiency of sampler specially in the case of substantially sophisticated target distribution.
- Also, PSAMH and PT in all cases delivered close results, however, PSAMH outweighed when the computational running time appears as an important issue.
- Moreover, bear in mind that, the aforementioned experiment by no mean has not disqualified other samplers, whereas it is known that, characteristics and complexity of introduced problem, initiate the justifiability of one sampler over the other.

- Furthermore, it should be noted that by a slight modification on each one of exerted samplers, it is possible to improve the performance of them which is thoroughly out of the scope of this study.

8.1.3. Section 4

- Researches employed PSAMH which is an Adaptive MCMC framework to calibrate two well-known empirical decline curve models (MHD and PLED). Applying PSAMH, the posterior space of model parameters as well as the event space are retrieved.
- To validate the persistency of results, eight experimental cases comprising four wells of the Eagle Ford Shale are examined. Exploiting the model realizations for the current and 30 years of production, several comparison plots of MHD and PLED are derived.
- In general, model parameters of MHD approximates the posterior space similar to the Normal distribution, while in contrary PLED random variables produced arbitrary distributions with irregular configurations.
- MHD manifested a consistent behavior regarding the current and 30 years' production and level of confidence in all experiments.
- MHD progressively developed more production with less confidence in compare to PLED.
- Eventually, it is implied that MHD overestimates or PLED underestimates the production. Also, the comparison plots are depicted the quick decline pattern in PLED which justifies the level of confidence in the production.

8.1.4. Section 5

- Employing various assimilation of forward model's parameters, alludes different statistics inferences.
- Increasing the size of random variables most often induces more uncertainty in the analysis.
- Although the order of confidence is varied from case to case, the overall magnitude of uncertainty is substantially small in compare to the production rate. Therefore, it justifies applying the Bayesian analysis in the case of calibration and model uncertainty reduction.
- The analogous of mean of realization amid various experimental designs and expert model, implies the underestimation of expert model in both the current and 20 years' production.
- The case of C_s connotes more confidence in compare to other experiments for short and long-term depletions.
- The correlation structure among the variables is irrefutably influenced by the variation of incorporated parameters (e.g. considering the same pair of parameters ($K_D^* - s$), from the positive quantity in one case ($C_{K_D^* - s}$) to negative value in the other case ($C_{K^0 - K_D^* - s}$)).

8.1.5. Section 6

- The evolution plots of the posterior distribution coupled with the MCMC experiment of random variable b elucidate the unreliability of early time production.
- The expected value of parameter b is inevitably correlated to the depletion time and can be dramatically altered by progressing in time.
- Although, the plots of sequential time intervals exhibit an extreme variation in the level of confidence in analysis by progressing in time, it is evident that the successful updating is highly correlated to the quality of observed data and the correlation structure between model random variables.
- The independent development in the location and size of updating posterior distribution, authenticates the functionality of hybrid prior mechanism. It should be taken into consideration that, if instead of the hybrid prior, merely the vector of former step posterior was employed, the current step posterior, under influence of constraint prior, would adhere to one region of random field. Hence, the concluded posterior could not appear as a true representative of the posterior space.
- Augmentation of observed data continuously improves the state of inference, which can be interpreted from the mean of realization plots. That is for, by increasing the time intervals, for example after 500 days, the layout of realization means curvatures almost undistinguishable.

- Also, the plots of mean and standard deviation of entire experimental designs associated with the short and long terms of production, irrefutably validate the advantage of exploiting the updating over regular method.

8.1.6. Section 7

- Providing a set of observed data, a precise forward model and corresponding locations would be enough to drive a spatial analysis.
- MCMC augmented with PSAMH framework allow us to efficiently and accurately sample from the random field.
- Extrapolation of forward model realizations in the form of daily basis or cumulative production can be used to feature dynamic Bayesian-spatial maps associated with varied time intervals.
- Exploiting the Bayesian analysis posterior, it is possible to delineate the forward model parameters regional map.
- The IDW, OK-Spherical, OK-Exponential, OK-Gaussian and OK-Stable spatial analysis techniques are employed to assess the efficiency of target location estimations. It is eventually implied that the OK-Exponential model developed a more precise spatial approximation tool comparing to other methods in both PLED and MHD empirical models.
- Clustering of locations, substantially increases the effectiveness of spatial analysis regarding PLED and MHD functions.

- A comparison amidst the aggregation of standard errors respected with the daily basis and cumulative productions connoted the undeniable advantage of MHD method according to the given production data.

8.2. Future research

According to the results obtained in this study, the subsequent future research is proposed:

- The batch size in PSAMH method which aims to identify the iteration number regarding the scaling factor, is set deterministically and by the course of several trial and errors. A new systematic approach can be generated to automatically assign an appropriate batch size corresponding to the convergence feature of MCMC.
- The number of asynchronous chains in PSAMH method becomes crucial when selecting the large number of chains makes the sampler inefficient or small number of chains decreases the likelihood of capturing the optimum step size. Therefore, it is relevant to run a research to optimize the number of chains automatically. This mechanism should be able to decrease or increase the number of chains before approaching to the stationary condition.
- The current research only takes the unconventional reservoir data into account. Hence, the performance of PSAMH sampler and entire provided results are under influence of hydraulic-fractured wells' behavior. It would be considerably beneficial, to evaluate the performance of the PSAMH sampler on the conventional reservoir data.

- There are varieties of analytical or semi-analytical reservoir modeling, which in this study only one of them is exerted. It would be valuable to evaluate the performance of other models exploiting PSAMH framework.
- The Bayesian inferences retrieved from 7.6.1Task 1of section 7, are obtained from scattered well locations associated with an unconventional development. It is evident that; unconventional reservoirs due to their individual characteristics and the employed depletion's technology are typically independent from the other neighbor wells. Thus, often it is uncertain that attaining information from one well could be accountable enough to expand it to other neighbor wells. Therefore, it is worthy to utilize the aforementioned spatial mechanism over a conventional reservoir.
- Considering the attribute of spatial analysis techniques, the precision of results is highly correlated to the distance amidst coordinates coupled with the regional density of known locations. Rerunning the provided algorithm in section 7 with a set of large number of wells in a smaller region will undeniably deliver a novel insight regarding the capabilities of proposed Bayesian-spatial approach.
- PSAMH is a well stablished method subordinates of the Bayesian paradigm that allows sampling from the complicated random fields. That is for, it becomes possible to exert the PSAMH sampling on varieties of other engineering applications such as Civil, Electrical or Mechanical Engineering when Bayesian analysis appears as a practical solution.

REFERENCES

Abdollahzadeh, Asaad, Alan Reynolds, Michael Mike Christie, David Corne, Brian Davies, and Glyn Williams. 2011. "Bayesian Optimization Algorithm Applied to Uncertainty Quantification." *Proceedings of SPE EUROPEC/EAGE Annual Conference and Exhibition*, no. September: 865–73. doi:10.2118/143290-MS.

Al-mudhafar, Watheq. 2015. "Integrating Bayesian Model Averaging for Uncertainty Reduction in Permeability Modeling." In *Offshore Technology Conference*, 1–20. Houston.

Al-mudhafar, Watheq J, William G A L Silva, and Coppe Ufrj. 2015. "Comparative Geostatistical Simulation of Formation Permeability through Bayesian and Conventional Kriging Approaches." In *Offshore Technology Conference*, 1–16. Rio de Janerio.

Andrieu, Christophe, Arnaud Doucet, and Roman Holenstein. 2010. "Particle Markov Chain Monte Carlo Methods." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 72 (3): 269–342. doi:10.1111/j.1467-9868.2009.00736.x.

Andrieu, Christophe, Nando De Freitas, and Arnaud Doucet. 1999. "Sequential MCMC for Bayesian Model Selection." *IEEE Journal of Higher-Order Statistics*.

Andrieu, Christophe, and Johannes Thoms. 2008. "A Tutorial on Adaptive MCMC." *Statistics and Computing* 18 (4): 343–73. doi:10.1007/s11222-008-9110-y.

Arps, J.J. 1944. "Analysis of Decline Curves." *Houston Meeting*, 20.

Arulampalam, M. Sanjeev, Simon Maskell, Neil Gordon, and Tim Clapp. 2002. "A

Tutorial on Particle Filters for Online Nonlinear/non-Gaussian Bayesian Tracking." *IEEE Transactions on Signal Processing* 50 (2): 174–88. doi:10.1109/78.978374.

Atchadé, Y, Gersende Fort, Eric Moulines, and Pierre Priouret. 2009. "Adaptive Markov Chain Monte Carlo : Theory and Methods." *Preprint*, 1–31. doi:10.1.1.192.2541.

Bachmaier, Martin, and Matthias Backes. 2008. "Variogram or Semivariogram? Understanding the Variances in a Variogram." *Precision Agriculture* 9 (3): 173–75. doi:10.1007/s11119-008-9056-2.

Banerjee, Sudipto, and Montserrat Fuentes. 2012. "Bayesian Modeling for Large Spatial Datasets." *Wiley Interdisciplinary Reviews: Computational Statistics* 4 (1): 59–66. doi:10.1002/wics.187.

Bayes, F.R.S, and Price. 1763. "An Essay towards Solving a Problem in the Doctrine of Chances." *The Royal Society* 53 (1763): 370–418.

Bédard, Mylène. 2007. "Weak Convergence of Metropolis Algorithms for Non-I.I.D. Target Distributions." *Annals of Applied Probability* 17 (December): 1222–44. doi:10.1214/105051607000000096.

Bedard, Mylene, and Jeffrey S Rosenthal. 2008. "Optimal Scaling of Metropolis Algorithms: Heading toward General Target Distributions." *Canadian Journal of Statistics* 34 (4): 1–19. doi:10.1002/cjs.5550360401.

Bedi, A, and J P Harrison. 2013. "A Comparison of Bayesian Techniques and Non-Probabilistic Models in Rock Engineering Design." *47th US Rock Mechanics / Geomechanics Symposium*, no. ARMA-2013-550: 10. doi:ARMA-2013-550.

Berg, Bernd A., and Alian Billoire. 2008. "Markov Chain Monte Carlo Simulations." In

Advanced Markov Chain Monte Carlo Methods, 1–14. John Wiley & Sons, Inc.

Bohling, Geoff. 2005a. "INTRODUCTION TO GEOSTATISTICS And VARIOGRAM ANALYSIS." *Earth*, no. October: 1–20. doi:10.1162/0162287054769931.

———. 2005b. "Kriging." *Kansas Geological Survey*, no. October: 1–20. doi:10.2104/ag050010.

Box, George E. P., Gwilym M. Jenkins, and Gregory C. Reinsel. 2008. *Time Series Analysis, Forcasting and Control*. Fourht. WILEY.

Brown, P C, and G K Falade. 2003. "A Quick Look Kriging Technique for Reservoir Charactersiation." *SPE International Technical Conference and Exhibition*.

Burkardt, John. 2014. "The Truncated Normal Distribution." *Department of Scientific Computing Website, Florida State University*. http://people.sc.fsu.edu/~jburkardt/presentations/truncated.

Calderhead, Ben. 2014. "A General Construction for Parallelizing Metropolis-Hastings Algorithms." *Proceedings of the National Academy of Sciences of the United States of America* 111 (49): 17408–13. doi:10.1073/pnas.1408184111.

Capen, E C. 1976. "The Difficulty of Assessing Uncertainty." *J. Petroleum Technology* 28 (8): 843–50. doi:10.2118/5579-PA.4.

Caruso, C., and F. Quarta. 1998. "Interpolation Methods Comparison." *Computers & Mathematics with Applications* 35 (12): 109–26. doi:10.1016/S0898-1221(98)00101-1.

Carvalho, Carlos M., Michael S. Johannes, Hedibert F. Lopes, and Nicholas G. Polson. 2010. "Particle Learning and Smoothing." *Statistical Science* 25 (1): 88–106.

doi:10.1214/10-STS325.

Chen, Rong, and Jun S Liu. 2000. “Mixture Kalman Filters.” *J.R. Statist. Soc.B* 62 (3): 493–508. doi:10.1111/1467-9868.00246.

Chen, Z H E. 2003. “Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond.” *Statistics* 182 (1): 1–69. doi:10.1.1.107.7415.

Cheng, Yueming, Yuhong Wang, Duane McVay, and W. John Lee. 2010. “Practical Application of a Probabilistic Approach to Estimate Reserves Using Production Decline Data.” *SPE Economics & Management* 2 (1): 9–12. doi:10.2118/95974-PA.

Cipra, By Barry A. 2000. “The Best of the 20th Century : Editors Name Top 10 Algorithms.” *SIAM News* 33 (4): 20–21.

Cotter, Colin, Simon Cotter, and Paul Russell. 2015. “Parallel Adaptive Importance Sampling.” *arXiv Preprint arXiv,* 1–21. http://arxiv.org/abs/1508.01132.

Craiu, Radu V., Lawrence Gray, Krzysztof Łatuszyński, Neal Madras, Gareth O. Roberts, and Jeffrey S. Rosenthal. 2014. “Stability of Adversarial Markov Chains, with an Application to Adaptive MCMC Algorithms.” *Annals of Applied Probability* 25 (6): 3592–3623. doi:10.1214/14-AAP1083.

Craiu, Radu V., and Christiane Lemieux. 2007. “Acceleration of the Multiple-Try Metropolis Algorithm Using Antithetic and Stratified Sampling.” *Statistics and Computing* 17 (2): 109–20. doi:10.1007/s11222-006-9009-4.

Craiu, Radu V., Jeffrey Rosenthal, and Chao Yang. 2009. “Learn From Thy Neighbor: Parallel-Chain and Regional Adaptive MCMC.” *Journal of the American Statistical Association* 104 (488): 1454–66. doi:10.1198/jasa.2009.tm08393.

Deutsch, C. V., and S. D. Zanon. 2007. “Direct Prediction of Reservoir Performance with Bayesian Updating.” *Journal of Canadian Petroleum Technology* 46 (2): 22–26. doi:10.2118/07-02-02.

Doucet, Arnaud, Simon Godsill, and Christophe Andrieu. 2000. “On Sequential Monte Carlo Sampling Methods for Bayesian Filterin.” *Statistics and Computing*, 197–208. doi:10.1023/A:1008935410038.

Earl, David J., and Michael W. Deem. 2005. “Parallel Tempering: Theory, Applications, and New Perspectives.” *Physical Chemistry Chemical Physics* 7 (23): 3910–16. doi:10.1039/b509983h.

Elsheikh, A. H., M. D. Jackson, and T. C. Laforce. 2012. “Bayesian Reservoir History Matching Considering Model and Parameter Uncertainties.” *Mathematical Geosciences* 44 (5): 515–43. doi:10.1007/s11004-012-9397-2.

Everitt, B.S. 2006. *The Camberidge Dictionary of Statistics*. Third. Cambridge: Cambridge University Press.

Everitt, B.S., and A. Skrondal. 2010. *The Cambridge Dictionary of Statistics*. Forth. New York: Cambridge University Press.

Faming, Liang, Liu Chuanhai, and Carroll Raymond. 2010. *Advanced Markov Chain Monte Carlo Methods: Learning from Past Samples*. John Wiley & Sons.

Fetkovich, M.J., E.J. Fetkovich, and M.D. Fetkovich. 1996. “Useful Concepts for Decline Curve Forecasting, Reserve Estimation, and Analysis.” *SPE Reservoir Engineering* 11 (1). doi:10.2118/28628-PA.

Floris, F. J. T., M. D. Bush, M. Cuypers, F. Roggero, and a.-R. Syversveen. 2001.

"Methods for Quantifying the Uncertainty of Production Forecasts: A Comparative Study." *Petroleum Geoscience* 7 (S): S87–96. doi:10.1144/petgeo.7.S.S87.

Forkman, Johannes. 2009. "Estimator and Tests for Common Coefficients of Variation in Normal Distributions." *Communications in Statistics-Theory and Methods* 38 (2): 233–51. http://dx.doi.org/10.1080/03610920802187448%0AAccess.

Forkman, Johannes, and Steve Verrill. 2008. "The Distribution of McKay ' S Approximation for the Coefficient of Variation." *Statistics and Probability Letters* 78 (1): 10–14. http://dx.doi.org/10.1016/j.spl.2007.04.018.

Fotheringham, a. S., M. E. Charlton, and C. Brunsdon. 1998. "Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis." *Environment and Planning A* 30 (11): 1905–27. doi:10.1068/a301905.

Fuentes-Cruz, G., E. Gildin, and P.P. Valkó. 2014. "On the Analysis of Production Data : Practical Approaches for Hydraulically Fractured Wells in Unconventional Reservoirs." *SPE Hydraulic Fracturing Technology Conference*.

Fuentes-cruz, Gorgonio, Eduardo Gildin, and Peter P Valkó. 2013. "Analyzing Production Data From Hydraulically Fractured Wells : The Concept of Induced Permeability Field." *Spe 163843*, no. 1994. doi:10.2118/163843-MS.

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. 2013. *Bayesian Data Analysis*. Third. Chapman & Hall/CRC.

Gelman, Andrew, Xiao-Li Meng, and Hal Stern. 1996. "Posterior Predictive Assessment of Model Fitness via Realized Discrepancies." *Statistica Sinica* 6 (4): 733–807. doi:10.1.1.142.9951.

Gelman, Andrew, G Roberts, and Wr Gilks. 1996. “Efficient Metropolis Jumping Rules.” *Bayesian Statistics*. Oxford university press. http://www.stat.columbia.edu/~gelman/research/published/baystat5.pdf.

Geyer, Charles J. 2002. “Introduction to Markov Chain Monte Carlo.” *Handbook of Markov Chain Monte Carlo,* no. 1990: 3–48. http://www.mcmchandbook.net/index.html.

Gilks, W.R, and S. Richardson. 1996. *Markov Chain Monte Carlo in Practice: Interdisciplinary Statistics*. First Edit. Chapman & Hall/CRC.

Gong, Xinglai, Raul Gonzalez, D a Mcvay, and Jeff Hart. 2011. “Bayesian Probabilistic Decline Curve Analysis Quantifies Shale Gas Reserves Uncertainty.” *Canadian Unconventional Resources & International Petroleum Conference,* no. November: 1–8.

Gong, Xinglai, Raul Gonzalez, Duane a Mcvay, Jeffrey D Hart, and a Texas. 2014. “Bayesian Probabilistic Decline-Curve Analysis Reliably Quantifies Uncertainty in Shale-Well-Production Forecasts,” no. December. doi:10.2118/147588-PA.

Gotway, C A, R B Ferguson, G W Hergert, and T A Peterson. 1996. “Comparison of Kriging and Inverse-Distance Methods for Mapping Soil Parameters.” *Soil Science Society of America Journal* 60 (4): 1237–47. doi:10.2136/sssaj1996.03615995006000040040x.

Graves, Todd L. 2011. “Automatic Step Size Selection in Random Walk Metropolis Algorithms.” *arXiv Preprint arXiv,* 1–13. http://arxiv.org/abs/1103.5986.

Haario, Heikki, Eero Saksman, and Johanna Tamminen. 1999. “Adaptive Proposal

Distribution for Random Walk Metropolis Algorithm." *Computational Statistics* 14 (3): 375. doi:10.1007/s001800050022.

———. 2001. "An Adaptive Metropolis Algorithm." *Bernoulli* 7 (2): 223–42. doi:10.2307/3318737.

Hasting, W K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrica* 57 (1): 97.

Hastings, W K. 1970. "Monte Carlo Sampling Methods Using Markov Chain and Their Applications." *Biometrika*.

Held, Leonhard, Brigit Schrodle, and Harvard Rue. 2010. "Posterior and Cross-Validatory Predictive Checks: A Comparison of MCMC and INLA." *Statistical Modelling and Regression Structures: Festschrift in Honour of Ludwig Fahrmeir*, no. January 2010: 1–472. doi:10.1007/978-3-7908-2413-1.

Hoff, Peter D. 2009. *A First Course in Bayesian Statistical Methods*. New York: Springer.

Hukushima, K., and K. Nemoto. 1996. "Exchange Monte Carlo Method and Application to Spin Glass Simulations." *Journal of the Physical Society of Japan*. doi:10.1143/JPSJ.65.1604.

Ibegbuna, B, W Brent, William M Cobb, and C Marshall. 2012. "Assessment of Probabilistic Parameters for Barnett Shale Recoverable Volumes." In *SPE Hydrocarbon Economics and Evaluation Symposium*, 16. Calgary.

Ilk, D., a.D. Perego, J.a. Rushing, and T.a. Blasingame. 2008. "Exponential vs. Hyperbolic Decline in Tight Gas Sands — Understanding the Origin and

Implications for Reserve Estimates Using Arps' Decline Curves." *Spe-116731* 116731. doi:10.2118/116731-MS.

Isaaks, E H, and R M Srivastava. 1989. *Applied Geostatistics*. *New York*. Vol. 21. doi:10.1016/0040-1951(74)90006-7.

Jimenez, E A, E A Idrobo, J Ernandez, and R Startzman. 2005. "Optimizing Estimates of Probabilistic Reserves from Production Trends Using a Bayesian Approach." In *Society of Petroleum Engineers*. Dallas.

Jochen, V A, J P Spivey, and S A Holditch. 1996. "Probabilistic Reserves Estimation Using Decline Curve Analysis With the Bootstrap Method." *Society of Petroleum Engineers*, 589–96.

Kendall, W.S, F Liang, and J-S Wang. 2005a. *Markov Chain Monte Carlo: Innovations and Applications*. *Mathematical Biosciences and Engineering : MBE*. Vol. 3. http://www.ncbi.nlm.nih.gov/pubmed/21157593.

Kendall, W S, F Liang, and J-S Wang. 2005b. *MARKOV CHAIN MONTE CARLO Inovations and Applications*. World Scientific Publishing Co.

Lauritzen, Steffen. 2008. "Sequential Bayesian Updating," 1–21.

Lee, Dominic S., and Nicholas K K Chia. 2002. "A Particle Algorithm for Sequential Bayesian Parameter Estimation and Model Selection." *IEEE Transactions on Signal Processing* 50 (2): 326–36. doi:10.1109/78.978387.

Lee, W. John, and Rod Sidle. 2010. "Gas-Reserves Estimation in Resource Plays." *SPE Economics & Management* 2 (2): 23–25. doi:10.2118/130102-PA.

Lehmann, a, J M Overton, and J R Leathwick. 2002. "GRASP: Generalized Regression

Analysis and Spatial Prediction." *Ecological Modelling* 157 (2–3): 189–207. doi:10.1016/S0304-3800(02)00195-3.

Liu, Jun S, Faming Liang, and Wing Hung Wong. 2000. "The Multiple-Try Method and Local Optimization in Metropolis Sampling." *Journal of the American Statistical Association* 95 (449): 121–34. doi:10.2307/2669532.

Lopes, Hedibert F., Michael S. Johannes, Carlos M. Carvalho, Nicholas G. Polson, and Michael Pitt. 2012. "Particle Learning for Sequential Bayesian Computation." *Bayesian Statistics 9*. doi:10.1093/acprof:oso/9780199694587.003.0011.

Lu, George Y., and David W. Wong. 2008. "An Adaptive Inverse-Distance Weighting Spatial Interpolation Technique." *Computers and Geosciences* 34 (9): 1044–55. doi:10.1016/j.cageo.2007.07.010.

Mattar, L, and D M Anderson. 2003. "A Systematic and Comprehensive Methodology for Advanced Analysis of Production Data." *Proceedings - SPE Annual Technical Conference and Exhibition*, 3723. doi:10.2118/84472-MS.

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. "Equation of State Calculations by Fast Computing Machines." *The Journal of Chemical Physics* 21 (6): 1087–92. doi:doi:10.1063/1.1699114.

Miasojedow, Błażej, Eric Moulines, and Matti Vihola. 2013. "An Adaptive Parallel Tempering Algorithm." *Journal of Computational and Graphical Statistics* 22 (3): 1–33. doi:10.1080/10618600.2013.778779.

Mitas, L, and H Mitasova. 1999. "Spatial Interpolation." *Geographical Information*

Systems: Principles, Techniques, Management and Applications, 481–92. http://skagit.meas.ncsu.edu/~helena/gmslab/papers/hgint39.pdf.

Moridis, Nefeli, Yasser Soltanpour, Zenon Medina-Cetina, W. John Lee, and Thomas A. Blasingame. 2017. "A Production Characterization of the Eagle Ford Shale, Texas - A Bayesian Analysis Approach." doi:10.2118/185597-MS.

Purvis, D C, and Heidi Kuzma. 2016. "Evolution of Uncertainty Methods in Decline Curve Analysis." *Society of Petroleum Engineers*, no. May: 17–18.

Pyrcz, Michael J., and C. V. Deutsch. 2014. *GEOSTATISTICAL RESERVOIR MODELING*. Second. New York: Oxford university press.

Robert, Christian P. 1995. "Simulation of Truncated Normal Variables." *Statistics and Computing* 5 (2): 121–25. doi:10.1007/BF00143942.

Roberts, G. O., A. Gelman, and W. R. Gilks. 1997. "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms." *Annals of Applied Probability* 7 (1): 110–20. doi:10.1214/aoap/1034625254.

Roberts, Gareth O., and J. S. Rosenthal. 2002. "Optimal Scaling of Various Metropolis-Hastings Algorithms." 16 (4): 351–67. doi:10.1214/ss/1015346320.

Roberts, Gareth O., and Jeffrey S. Rosenthal. 2009. "Examples of Adaptive MCMC." *Journal of Computational and Graphical Statistics* 18 (2): 349–67. doi:10.1198/jcgs.2009.06134.

Roberts, Gareth O, Jeffrey S Rosenthal, Gareth O R Oberts, and Jeffrey S R Osenthal. 2009. "Examples of Adaptive MCMC." *Journal of Computational Science and Graphical Statistics* 18:2: 349–67. doi:10.1198/jcgs.2009.06134.

Robertson, S. 1988. “Generalized Hyperbolic Equation.” In *Society of Petroleum Engineers*.

Roggero, Frédéric, and Dominique Guérillot. 1996. “Gradient Method and Bayesian Formalism Application to Petrophysical Parameter Characterization.” *Proceedings of the 5th European Conference on the Mathematics of Oil Recovery*, 271–82.

Rosenthal, Jeffrey S. 2011. “Optimal Proposal Distributions and Adaptive MCMC.” *Handbook of Markov Chain Monte Carlo*, no. 1: 93–112. doi:10.1201/b10905.

Rosenthal, Jeffrey S, S Brooks, A Gelman, and G Jones. 2010. “Optimal Proposal Distributions and Adaptive MCMC,” no. 1: 1–25.

Rotondi, Marco, Giovanna Nicotra, Antonella Godi, F. Michela Contento, Martin Blunt, and Mike Christie. 2006. “Hydrocarbon Production Forecast and Uncertainty Quantification: A Field\nApplication.” *Proceedings of SPE Annual Technical Conference and Exhibition*, 1–9. doi:10.2523/102135-MS.

Sarkka, Simo. 2013. “Bayesian Filtering and Smoothing.” *Cambridge University Press*, 254. doi:10.1017/CBO9781139344203.

Septier, Fran??ois, and Gareth W. Peters. 2016. “Langevin and Hamiltonian Based Sequential MCMC for Efficient Bayesian Filtering in High-Dimensional Spaces.” *IEEE Journal on Selected Topics in Signal Processing* 10 (2): 312–27. doi:10.1109/JSTSP.2015.2497211.

Seshadri, J, and L Mattar. 2010. “Comparison of Power Law and Modified Hyperbolic Decline Methods.” *Canadian Unconventional Resources and International Petroleum Conference*, no. October: 19–21.

Sidler, Rolf, Supervisor Prof, and Klaus Holliger. 2003. “Kriging and Conditional Geostatistical Simulation Based on Scale-Invariant Covariance Models by.” *Unpublished Diploma Thesis. Institute of Geophysics, Department of Earth Science, Swiss Federal Institute of Technology Zurich*, no. March: 79. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.126.3499&rep=rep1&type=pdf.

Sims, Christopher A. 1998. “Adaptive Metropolis-Hasting Sampling, or Monte Carlo Kernel Estimation.” *University of Princeton*.

Tarrahi, Mohammadali, Sergio Gonzales, and Eduardo Gildin. 2014. “Real-Time Estimation of Hydraulic Fracture Characteristics from Production Data.” *Proceedings of the 2nd Unconventional Resources Technology Conference*. doi:10.15530/urtec-2014-1923687.

Tarrahi, Mohammadali, Shell Global Solutions, Sardar Afra, and A Texas. 2016. “Markov-Chain Monte Carlo with Locally Varying Mean Kriging for Improved Reservoir Model Calibration and Uncertainty Assessment.”

Thompson, Madeleine B. 2010a. “A Comparison of Methods for Computing Autocorrelation Time.” *ArXiv E-Prints*, no. 1007: 8. http://arxiv.org/abs/1011.0175.

———. 2010b. “A Comparison of Methods for Computing Autocorrelation Time.” *ArXiv E-Prints*, no. 1007: 8.

Tian, Yao, Walter B Ayers, Huiyan Sang, and William D Mccain. 2017. “Quantitative Evaluation of Key Geological Controls on Regional Eagle Ford Shale Production Using Spatial Statistics.” *SPE Unconventional Resources Conference*.

Van Brummelen, Glen. 2012. *Heavenly Mathematics : The Forgotten Art of Spherical Trigonometry*. Princeton University Press.

Vega, Leonardo, Danny Rojas, and Akhil Datta-Gupta. 2004. "Scalability of the Deterministic and Bayesian Approaches to Production-Data Integration Into Reservoir Models." *SPE Journal* 9 (3): 3–5. doi:10.2118/88961-PA.

Vidakovic, Brani. 1962. "Bayesian Statistics." *Bayesian Statistics for Engineers* 1: 1–22. http://zoe.bme.gatech.edu/~bv20/bmestat/Bank/bayes.pdf.

Vink, Jeroen C, and Guohua Gao. 2015. "Bayesian Style History Matching: Another Way to Under-Estimate Forecast Uncertainty?" *SPE Annual Technical Conference and Exhibition*, 28–30. doi:10.2118/175121-MS.

Wackemagel, Hans. 2003. *Multivariate Geostatistics An Introduction with Applications*. Springer.

Wang, Jian-Sheng, and Robert H Swendsen. 2004. "Replica Monte Carlo Simulation (Revisited)." *Exchange Organizational Behavior Teaching Journal* 157 (1): 7. doi:10.1143/PTPS.157.317.

Weber, Dennis, and Evan Englund. 1992. "Evaluation and Comparison of Spatial Interpolators." *Mathematical Geology* 24 (4): 381–91. doi:10.1007/BF00891270.

Whiteley, Nick, Christophe Andrieu, and Arnaud Doucet. 2010. "Efficient Bayesian Inference for Switching State-Space Models Using Discrete Particle Markov Chain Monte Carlo Methods," 1–26. http://arxiv.org/abs/1011.2437.

Willigers, Bart J.A., Steve Begg, and Reidar B. Bratvold. 2014. "Combining Geostatistics With Bayesian Updating To Continually Optimize Drilling Strategy in

Shale-Gas Plays." *SPE Reservoir Evaluation & Engineering* 17 (4): 507–19. doi:10.2118/164816-PA.

Xie, Jiang, Yalchin Efendiev, and Akhil Datta-Gupta. 2011. "Uncertainty Quantification in History Matching of Channelized Reservoirs Using Markov Chain Level Set Approaches." *SPE Reservoir Simulation Symposium*. doi:10.2118/141811-MS.

Xifara, T., C. Sherlock, S. Livingstone, S. Byrne, and M. Girolami. 2014. "Langevin Diffusions and the Metropolis-Adjusted Langevin Algorithm." *Statistics and Probability Letters* 91 (1): 14–19. doi:10.1016/j.spl.2014.04.002.

Yang, Jinyoung, Radu V. Craiu, and Jeffrey S. Rosenthal. 2016. "Adaptive Component-Wise Multiple-Try Metropolis Sampling." *arXiv Preprint arXiv*, 1–36. http://arxiv.org/abs/1603.03510.

Yang, Yun, and David B Dunson. 2013. "Sequential Markov Chain Monte Carlo." *arXiv.org* math.ST: 1–30. http://arxiv.org/abs/1308.3861v1%5Cnpapers2://publication/uuid/AD6545CF-A519-496A-AFA3-6F6D7C5D276B.

Yasrebi, Jafar, Mahboub Saffari, Hamed Fathi, and Najafali Karimian. 2009. "Evaluation and Comparison of Ordinary Kriging and Inverse Distance Weighting Methods for Prediction of Spatial Variability of Some Chemical Parameters." *Research Journal of Biological Science*.

Zhang, Guangzhi, Yongshe Liu, and Xingyao Yin. 2005. "The Effects of Factors of Spatial Correlation Analysis On Kriging Estimate in Geostatistical Reservoir Characterization," no. 3: 1457–61.

Zhang, Ye. 2009. *Introduction to Geostatistics - Course Notes*. *Course Notes*. papers2://publication/uuid/20990432-8B19-4169-AC3A-F73E130F2D04.

Zhang, Yunxiang, and Sanjay Srinivasan. 2005. “Markov Chain Monte Carlo for Modelling Permeability Variations in Reservoirs.” *Journal of Canadian Petroleum Technology* 44 (3): 40–45. doi:10.2118/05-03-03.

APPENDIX A: TASK 1, BAYESIAN ANALYSIS RESULTS

A.1. MHD

A.1.1. MHD model parameters

Table A.1 Mean and standard deviation of MHD parameters

Parameters	b		D_i		q_i	
Name	μ	σ	μ	σ	μ	σ
Well 1	4.42E-04	8.86E-04	2.05E-03	3.56E-05	1.15E+03	7.85E+00
Well 2	3.15E-04	3.57E-04	2.35E-03	2.07E-05	8.28E+02	3.16E+00
Well 3	1.88E+00	1.13E-01	1.96E-03	1.55E-04	4.90E+02	6.84E+00
Well 4	6.48E-04	1.59E-03	6.71E-03	1.43E-04	7.05E+02	1.05E+01
Well 5	1.67E+00	2.14E-01	4.80E-03	5.73E-04	3.38E+02	8.13E+00
Well 6	3.94E-04	6.45E-04	3.48E-03	5.11E-05	6.91E+02	5.19E+00
Well 7	1.00E-02	8.50E-02	5.83E-04	5.91E-05	3.16E+02	3.53E+00
Well 8	1.89E+00	1.01E-01	1.64E-03	9.57E-05	3.32E+02	3.10E+00
Well 9	1.26E-03	9.39E-03	8.30E-04	3.94E-05	3.58E+02	2.74E+00
Well 10	3.89E-04	6.35E-04	5.82E-03	5.98E-05	9.00E+02	5.78E+00
Well 11	1.66E-01	3.06E-02	5.77E-03	1.36E-04	1.04E+03	6.80E+00
Well 12	4.36E-04	8.75E-04	4.51E-03	3.79E-05	8.28E+02	4.00E+00
Well 13	8.24E-01	1.65E-01	2.96E-03	2.13E-04	1.01E+03	1.05E+01
Well 14	1.40E+00	1.27E-01	4.97E-03	3.47E-04	6.91E+02	7.98E+00
Well 15	1.98E+00	2.00E-02	5.62E-03	2.19E-04	7.83E+02	8.38E+00
Well 16	4.85E-04	1.22E-03	2.16E-03	2.30E-05	5.01E+02	2.03E+00
Well 17	1.12E-03	6.58E-03	9.33E-04	4.22E-05	4.79E+02	3.92E+00
Well 18	1.94E+00	5.72E-02	3.67E-03	2.30E-04	6.36E+02	9.33E+00
Well 19	1.76E+00	2.17E-01	3.18E-03	3.06E-04	7.20E+02	1.14E+01
Well 20	6.44E-04	2.06E-03	2.31E-03	6.38E-05	6.96E+02	7.93E+00
Well 21	8.94E-01	3.09E-02	1.49E-02	6.56E-04	1.74E+03	2.67E+01
Well 22	1.97E+00	2.17E-02	4.83E-02	4.43E-03	2.48E+03	9.19E+01
Well 23	3.53E-01	9.61E-02	4.28E-03	3.03E-04	1.25E+03	2.47E+01
Well 24	1.00E-03	4.24E-03	1.75E-03	5.43E-05	5.27E+02	5.19E+00
Well 25	6.66E-04	2.28E-03	2.73E-03	4.70E-05	7.81E+02	6.11E+00
Well 26	4.21E-01	9.60E-02	6.16E-03	5.14E-04	1.09E+03	2.51E+01
Well 27	1.65E+00	2.01E-01	7.99E-03	1.14E-03	4.44E+02	1.42E+01
Well 28	2.19E-03	7.21E-03	1.57E-03	5.50E-05	2.74E+02	2.68E+00
Well 29	1.90E+00	8.72E-02	2.33E-03	1.21E-04	4.62E+02	4.40E+00
Well 30	4.67E-04	1.10E-03	2.43E-03	2.40E-05	4.67E+02	2.02E+00
Well 31	6.57E-04	2.21E-03	2.40E-03	6.00E-05	5.11E+02	5.05E+00
Well 32	2.08E-03	1.22E-02	4.65E-03	8.72E-05	7.68E+02	7.83E+00
Well 33	7.17E-04	3.29E-03	2.04E-03	4.17E-05	6.98E+02	5.28E+00
Well 34	1.51E-03	6.63E-03	1.28E-03	4.80E-05	1.02E+03	9.39E+00
Well 35	1.90E+00	8.58E-02	6.97E-03	8.54E-04	3.86E+02	1.35E+01
Well 36	1.98E+00	2.06E-02	3.82E-03	1.36E-04	8.67E+02	7.93E+00
Well 37	1.98E+00	1.81E-02	6.38E-03	1.88E-04	1.23E+03	1.06E+01
Well 38	3.64E-01	6.31E-02	4.84E-03	2.05E-04	1.05E+03	1.05E+01
Well 39	5.36E-04	1.37E-03	4.79E-03	5.00E-05	9.14E+02	6.31E+00
Well 40	6.87E-04	1.39E-03	5.11E-03	1.22E-04	1.52E+03	2.52E+01
Well 41	8.18E-04	1.59E-03	4.91E-03	8.25E-05	1.14E+03	1.26E+01
Well 42	4.73E-04	1.04E-03	3.20E-03	4.34E-05	9.97E+02	6.91E+00
Well 43	4.06E-04	6.63E-04	2.26E-03	2.98E-05	4.43E+02	3.32E+00

A.1.2. Expected values of MHD daily basis production

Table A.2 Mean of daily production associated with MHD model for several time intervals

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	5.50E+02	2.75E+01	6.68E-01	1.64E-02	4.08E-04	1.03E-05	2.67E-07
Well 2	3.56E+02	1.14E+01	1.58E-01	2.21E-03	3.11E-05	4.41E-07	6.32E-09
Well 3	3.13E+02	1.64E+02	9.70E+01	5.73E+01	3.38E+01	2.00E+01	1.18E+01
Well 4	6.35E+01	3.68E-03	3.97E-08	9.36E-11	1.88E-12	7.63E-14	4.98E-15
Well 5	1.51E+02	6.55E+01	3.85E+01	2.27E+01	1.34E+01	7.92E+00	4.68E+00
Well 6	1.98E+02	1.21E+00	2.18E-03	4.04E-06	7.78E-09	1.62E-11	4.29E-14
Well 7	2.57E+02	1.10E+02	3.91E+01	1.41E+01	5.22E+00	1.98E+00	7.75E-01
Well 8	2.24E+02	1.21E+02	7.14E+01	4.22E+01	2.49E+01	1.47E+01	8.69E+00
Well 9	2.66E+02	7.91E+01	1.76E+01	3.95E+00	8.98E-01	2.08E-01	4.98E-02
Well 10	1.11E+02	2.24E-02	5.98E-07	1.88E-11	1.85E-15	2.20E-18	7.61E-21
Well 11	1.75E+02	2.45E+00	1.46E-01	2.39E-02	6.42E-03	2.29E-03	9.88E-04
Well 12	1.66E+02	2.24E-01	6.30E-05	1.93E-08	1.16E-11	8.25E-14	1.91E-15
Well 13	4.70E+02	1.28E+02	6.24E+01	3.64E+01	2.15E+01	1.27E+01	7.50E+00
Well 14	2.83E+02	1.06E+02	6.15E+01	3.63E+01	2.14E+01	1.27E+01	7.48E+00
Well 15	3.48E+02	1.67E+02	9.84E+01	5.81E+01	3.43E+01	2.03E+01	1.20E+01
Well 16	2.31E+02	9.76E+00	1.92E-01	3.81E-03	7.65E-05	1.57E-06	3.35E-08
Well 17	3.42E+02	8.74E+01	1.61E+01	3.00E+00	5.65E-01	1.08E-01	2.13E-02
Well 18	3.31E+02	1.63E+02	9.60E+01	5.67E+01	3.35E+01	1.98E+01	1.17E+01
Well 19	3.85E+02	1.81E+02	1.07E+02	6.31E+01	3.72E+01	2.20E+01	1.30E+01
Well 20	3.03E+02	1.03E+01	1.58E-01	2.49E-03	4.39E-05	1.26E-06	1.18E-07
Well 21	2.45E+02	4.69E+01	2.21E+01	1.30E+01	7.69E+00	4.54E+00	2.68E+00
Well 22	4.09E+02	1.81E+02	1.07E+02	6.33E+01	3.74E+01	2.21E+01	1.30E+01
Well 23	3.66E+02	2.99E+01	6.82E+00	2.77E+00	1.44E+00	8.26E-01	4.85E-01
Well 24	2.82E+02	2.19E+01	9.31E-01	4.07E-02	1.88E-03	1.01E-04	8.59E-06
Well 25	2.93E+02	5.40E+00	3.84E-02	2.92E-04	3.22E-06	1.54E-07	2.39E-08
Well 26	2.29E+02	1.78E+01	4.60E+00	2.05E+00	1.13E+00	6.61E-01	3.90E-01
Well 27	1.55E+02	6.30E+01	3.70E+01	2.18E+01	1.29E+01	7.61E+00	4.49E+00
Well 28	1.56E+02	1.59E+01	9.54E-01	5.93E-02	3.95E-03	3.12E-04	3.77E-05
Well 29	2.80E+02	1.44E+02	8.52E+01	5.03E+01	2.97E+01	1.75E+01	1.04E+01
Well 30	1.95E+02	5.57E+00	6.72E-02	8.22E-04	1.03E-05	1.35E-07	2.08E-09
Well 31	2.17E+02	6.50E+00	8.48E-02	1.15E-03	1.76E-05	5.11E-07	5.71E-08
Well 32	1.45E+02	1.83E-01	9.49E-04	1.68E-04	5.41E-05	2.30E-05	1.16E-05
Well 33	3.35E+02	1.69E+01	4.15E-01	1.07E-02	3.89E-04	5.68E-05	2.40E-05
Well 34	6.42E+02	9.89E+01	9.78E+00	9.86E-01	1.03E-01	1.16E-02	1.65E-03
Well 35	1.54E+02	7.12E+01	4.20E+01	2.48E+01	1.46E+01	8.65E+00	5.11E+00
Well 36	4.47E+02	2.21E+02	1.31E+02	7.72E+01	4.56E+01	2.69E+01	1.59E+01
Well 37	5.18E+02	2.46E+02	1.45E+02	8.58E+01	5.07E+01	2.99E+01	1.77E+01
Well 38	2.74E+02	2.03E+01	4.47E+00	1.75E+00	8.89E-01	5.09E-01	2.99E-01
Well 39	1.65E+02	1.50E-01	2.64E-05	8.30E-09	5.51E-11	1.78E-12	9.53E-14
Well 40	2.43E+02	1.42E-01	1.52E-05	2.73E-09	9.61E-12	2.22E-13	8.74E-15
Well 41	1.96E+02	1.52E-01	2.28E-05	5.90E-09	1.95E-11	3.64E-13	1.23E-14
Well 42	3.16E+02	2.91E00	8.66E-03	2.65E-05	8.48E-08	2.99E-10	1.41E-12
Well 43	1.97E+02	7.18E+00	1.18E-01	1.94E-03	3.26E-05	5.54E-07	9.59E-09

A.1.3. Standard deviation of MHD daily basis production

Table A.3 Standard deviation of daily production regarding MHD model for 7 years

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	4.24E+00	1.64E+00	8.45E-02	3.29E-03	1.21E-04	4.92E-06	2.61E-07
Well 2	1.59E+00	3.87E-01	1.14E-02	2.47E-04	4.83E-06	9.02E-08	1.66E-09
Well 3	3.33E+00	5.06E+00	3.01E+00	1.78E+00	1.05E+00	6.19E-01	3.66E-01
Well 4	2.62E+00	1.17E-03	8.96E-07	7.44E-09	1.66E-10	7.08E-12	4.74E-13
Well 5	3.44E+00	6.51E+00	4.17E+00	2.46E+00	1.45E+00	8.59E-01	5.07E-01
Well 6	2.72E+00	1.08E-01	4.17E-04	1.32E-06	5.06E-09	3.53E-11	4.25E-13
Well 7	3.35E+00	1.10E+01	8.71E+00	5.26E+00	2.96E+00	1.66E+00	9.41E-01
Well 8	1.80E+00	2.92E+00	1.73E+00	1.02E+00	6.02E-01	3.55E-01	2.10E-01
Well 9	2.33E+00	5.30E+00	2.59E+00	1.03E+00	4.50E-01	2.33E-01	1.32E-01
Well 10	1.94E+00	2.50E-03	2.16E-07	6.42E-11	6.83E-14	1.58E-16	6.47E-19
Well 11	3.05E+00	8.05E-01	9.87E-02	2.29E-02	7.71E-03	3.25E-03	1.59E-03
Well 12	1.74E+00	1.57E-02	1.49E-05	4.70E-08	4.38E-10	7.23E-12	1.83E-13
Well 13	5.05E+00	2.01E+01	1.63E+01	1.02E+01	6.05E+00	3.58E+00	2.11E+00
Well 14	3.47E+00	8.08E+00	5.53E+00	3.27E+00	1.93E+00	1.14E+00	6.73E-01
Well 15	2.44E+00	1.87E+00	1.10E+00	6.51E-01	3.85E-01	2.27E-01	1.34E-01
Well 16	1.34E+00	3.83E-01	1.65E-02	5.64E-04	2.06E-05	1.00E-06	6.87E-08
Well 17	3.32E+00	6.22E+00	2.45E+00	7.30E-01	2.12E-01	7.11E-02	3.06E-02
Well 18	3.63E+00	3.75E+00	2.22E+00	1.31E+00	7.73E-01	4.56E-01	2.70E-01
Well 19	4.42E+00	1.38E+01	9.07E+00	5.37E+00	3.18E+00	1.88E+00	1.11E+00
Well 20	4.64E+00	1.12E+00	3.91E-02	1.72E-03	1.81E-04	3.02E-05	6.42E-06
Well 21	3.04E+00	2.66E+00	1.79E+00	1.07E+00	6.33E-01	3.74E-01	2.21E-01
Well 22	2.85E+00	2.39E+00	1.41E+00	8.34E-01	4.93E-01	2.91E-01	1.72E-01
Well 23	5.63E+00	9.56E+00	4.03E+00	2.16E+00	1.30E+00	7.78E-01	4.61E-01
Well 24	3.63E+00	2.02E+00	1.94E-01	1.74E-02	2.44E-03	5.47E-04	1.57E-04
Well 25	3.44E+00	4.52E-01	9.84E-03	4.81E-04	5.08E-05	7.97E-06	1.62E-06
Well 26	6.22E+00	6.18E+00	2.64E+00	1.48E+00	8.96E-01	5.34E-01	3.16E-01
Well 27	4.85E+00	7.39E+00	4.71E+00	2.78E+00	1.64E+00	9.71E-01	5.73E-01
Well 28	2.08E+00	1.50E+00	2.07E-01	2.75E-02	5.45E-03	1.65E-03	6.39E-04
Well 29	2.04E+00	3.47E+00	2.06E+00	1.21E+00	7.17E-01	4.23E-01	2.50E-01
Well 30	1.05E+00	2.28E-01	6.44E-03	1.61E-04	5.13E-06	2.35E-07	1.44E-08
Well 31	3.24E+00	6.66E-01	2.03E-02	8.57E-04	9.23E-05	1.54E-05	3.28E-06
Well 32	3.07E+00	1.96E-01	1.88E-02	4.85E-03	1.86E-03	8.87E-04	4.83E-04
Well 33	3.23E+00	1.25E+00	1.10E-01	2.52E-02	9.39E-03	4.25E-03	2.18E-03
Well 34	6.71E+00	8.02E+00	1.78E+00	3.51E-01	9.53E-02	3.75E-02	1.79E-02
Well 35	3.49E+00	3.10E+00	1.84E+00	1.09E+00	6.43E-01	3.80E-01	2.24E-01
Well 36	2.33E+00	2.15E+00	1.27E+00	7.51E-01	4.43E-01	2.62E-01	1.55E-01
Well 37	2.20E+00	2.03E+00	1.20E+00	7.09E-01	4.19E-01	2.47E-01	1.46E-01
Well 38	4.45E+00	4.93E+00	1.95E+00	9.95E-01	5.90E-01	3.58E-01	2.12E-01
Well 39	2.13E+00	1.50E-02	2.19E-05	1.83E-07	3.84E-09	1.44E-10	8.22E-12
Well 40	7.76E+00	3.08E-02	1.27E-05	5.75E-08	8.16E-10	2.12E-11	8.59E-13
Well 41	4.37E+00	2.36E-02	1.74E-05	7.54E-08	9.48E-10	2.21E-11	8.07E-13
Well 42	3.41E+00	2.16E-01	1.43E-03	8.25E-06	6.15E-08	7.19E-10	1.24E-11
Well 43	1.47E+00	3.58E-01	1.25E-02	3.29E-04	8.00E-06	1.98E-07	5.49E-09

A.1.4. Expected values of MHD cumulative production

Table A.4 Mean of cumulative production corresponding to MHD model

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	2.91E+05	5.47E+05	5.60E+05	5.60E+05	5.60E+05	5.60E+05	5.60E+05
Well 2	2.01E+05	3.48E+05	3.52E+05	3.52E+05	3.52E+05	3.52E+05	3.52E+05
Well 3	1.37E+05	4.58E+05	7.11E+05	9.07E+05	1.07E+06	1.22E+06	1.34E+06
Well 4	9.56E+04	1.05E+05	1.05E+05	1.05E+05	1.05E+05	1.05E+05	1.05E+05
Well 5	7.63E+04	2.14E+05	3.11E+05	3.82E+05	4.41E+05	4.92E+05	5.35E+05
Well 6	1.42E+05	1.98E+05	1.98E+05	1.98E+05	1.98E+05	1.98E+05	1.98E+05
Well 7	1.02E+05	3.58E+05	4.83E+05	5.28E+05	5.44E+05	5.50E+05	5.53E+05
Well 8	9.61E+04	3.30E+05	5.17E+05	6.63E+05	7.85E+05	8.93E+05	9.87E+05
Well 9	1.11E+05	3.38E+05	4.12E+05	4.28E+05	4.32E+05	4.33E+05	4.33E+05
Well 10	1.35E+05	1.55E+05	1.55E+05	1.55E+05	1.55E+05	1.55E+05	1.55E+05
Well 11	1.67E+05	2.14E+05	2.16E+05	2.16E+05	2.16E+05	2.16E+05	2.16E+05
Well 12	1.47E+05	1.84E+05	1.84E+05	1.84E+05	1.84E+05	1.84E+05	1.84E+05
Well 13	2.43E+05	5.89E+05	7.51E+05	8.43E+05	9.06E+05	9.53E+05	9.90E+05
Well 14	1.50E+05	3.88E+05	5.39E+05	6.43E+05	7.26E+05	7.95E+05	8.53E+05
Well 15	1.73E+05	5.06E+05	7.61E+05	9.57E+05	1.12E+06	1.27E+06	1.39E+06
Well 16	1.25E+05	2.28E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05
Well 17	1.46E+05	4.21E+05	4.97E+05	5.10E+05	5.13E+05	5.13E+05	5.14E+05
Well 18	1.56E+05	4.79E+05	7.28E+05	9.19E+05	1.08E+06	1.22E+06	1.34E+06
Well 19	1.82E+05	5.50E+05	8.23E+05	1.03E+06	1.20E+06	1.35E+06	1.47E+06
Well 20	1.70E+05	2.97E+05	3.02E+05	3.02E+05	3.02E+05	3.02E+05	3.02E+05
Well 21	2.07E+05	3.51E+05	4.09E+05	4.41E+05	4.63E+05	4.80E+05	4.92E+05
Well 22	2.53E+05	6.25E+05	9.00E+05	1.11E+06	1.29E+06	1.44E+06	1.58E+06
Well 23	2.47E+05	4.12E+05	4.38E+05	4.46E+05	4.49E+05	4.52E+05	4.53E+05
Well 24	1.41E+05	2.90E+05	3.02E+05	3.02E+05	3.02E+05	3.02E+05	3.02E+05
Well 25	1.79E+05	2.84E+05	2.86E+05	2.86E+05	2.86E+05	2.86E+05	2.86E+05
Well 26	1.82E+05	2.79E+05	2.95E+05	3.01E+05	3.04E+05	3.05E+05	3.06E+05
Well 27	8.49E+04	2.20E+05	3.13E+05	3.81E+05	4.36E+05	4.83E+05	5.24E+05
Well 28	7.52E+04	1.65E+05	1.75E+05	1.75E+05	1.75E+05	1.75E+05	1.75E+05
Well 29	1.25E+05	4.08E+05	6.29E+05	8.00E+05	9.43E+05	1.07E+06	1.18E+06
Well 30	1.12E+05	1.90E+05	1.92E+05	1.92E+05	1.92E+05	1.92E+05	1.92E+05
Well 31	1.23E+05	2.11E+05	2.13E+05	2.13E+05	2.13E+05	2.13E+05	2.13E+05
Well 32	1.34E+05	1.66E+05	1.66E+05	1.66E+05	1.66E+05	1.66E+05	1.66E+05
Well 33	1.78E+05	3.34E+05	3.42E+05	3.42E+05	3.42E+05	3.42E+05	3.42E+05
Well 34	2.93E+05	7.20E+05	7.90E+05	7.97E+05	7.98E+05	7.98E+05	7.98E+05
Well 35	7.97E+04	2.24E+05	3.32E+05	4.14E+05	4.83E+05	5.43E+05	5.95E+05
Well 36	2.12E+05	6.50E+05	9.90E+05	1.25E+06	1.47E+06	1.67E+06	1.84E+06
Well 37	2.62E+05	7.55E+05	1.13E+06	1.42E+06	1.66E+06	1.88E+06	2.07E+06
Well 38	1.95E+05	3.12E+05	3.30E+05	3.35E+05	3.37E+05	3.39E+05	3.39E+05
Well 39	1.56E+05	1.91E+05	1.91E+05	1.91E+05	1.91E+05	1.91E+05	1.91E+05
Well 40	2.49E+05	2.97E+05	2.97E+05	2.97E+05	2.97E+05	2.97E+05	2.97E+05
Well 41	1.92E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05
Well 42	2.13E+05	3.11E+05	3.11E+05	3.11E+05	3.11E+05	3.11E+05	3.11E+05
Well 43	1.09E+05	1.93E+05	1.96E+05	1.96E+05	1.96E+05	1.96E+05	1.96E+05

A.1.5. Standard deviation of MHD cumulative production

Table A.5 Standard deviation of cumulative production respected with MHD model

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	1.01E+03	5.86E+03	6.82E+03	6.86E+03	6.87E+03	6.87E+03	6.87E+03
Well 2	3.62E+02	1.85E+03	2.05E+03	2.05E+03	2.05E+03	2.05E+03	2.05E+03
Well 3	7.50E+02	6.94E+03	1.63E+04	2.58E+04	3.52E+04	4.44E+04	5.29E+04
Well 4	1.02E+03	1.50E+03	1.50E+03	1.50E+03	1.50E+03	1.50E+03	1.50E+03
Well 5	8.03E+02	8.69E+03	2.06E+04	3.20E+04	4.27E+04	5.27E+04	6.18E+04
Well 6	7.48E+02	2.02E+03	2.05E+03	2.05E+03	2.05E+03	2.05E+03	2.05E+03
Well 7	6.31E+02	1.30E+04	3.17E+04	4.44E+04	5.20E+04	5.66E+04	5.96E+04
Well 8	4.32E+02	4.06E+03	9.95E+03	1.61E+04	2.22E+04	2.81E+04	3.38E+04
Well 9	4.70E+02	7.33E+03	1.45E+04	1.76E+04	1.87E+04	1.92E+04	1.94E+04
Well 10	7.13E+02	1.12E+03	1.12E+03	1.12E+03	1.12E+03	1.12E+03	1.12E+03
Well 11	6.33E+02	3.66E+03	4.20E+03	4.28E+03	4.30E+03	4.31E+03	4.31E+03
Well 12	5.25E+02	1.10E+03	1.10E+03	1.10E+03	1.10E+03	1.10E+03	1.10E+03
Well 13	7.79E+02	2.39E+04	5.74E+04	8.50E+04	1.08E+05	1.28E+05	1.44E+05
Well 14	6.42E+02	1.04E+04	2.47E+04	3.79E+04	5.00E+04	6.12E+04	7.13E+04
Well 15	7.33E+02	3.64E+03	6.79E+03	9.60E+03	1.22E+04	1.47E+04	1.69E+04
Well 16	3.17E+02	1.73E+03	1.95E+03	1.96E+03	1.96E+03	1.96E+03	1.96E+03
Well 17	7.00E+02	9.19E+03	1.66E+04	1.91E+04	1.98E+04	1.99E+04	2.00E+04
Well 18	1.04E+03	6.35E+03	1.34E+04	2.01E+04	2.65E+04	3.27E+04	3.85E+04
Well 19	1.10E+03	1.63E+04	4.32E+04	7.00E+04	9.58E+04	1.20E+05	1.43E+05
Well 20	1.17E+03	5.51E+03	6.11E+03	6.14E+03	6.14E+03	6.14E+03	6.14E+03
Well 21	7.54E+02	5.19E+03	9.19E+03	1.21E+04	1.43E+04	1.62E+04	1.77E+04
Well 22	1.19E+03	4.47E+03	8.48E+03	1.22E+04	1.56E+04	1.88E+04	2.17E+04
Well 23	1.22E+03	1.84E+04	2.98E+04	3.50E+04	3.79E+04	3.98E+04	4.11E+04
Well 24	8.09E+02	5.66E+03	7.05E+03	7.16E+03	7.17E+03	7.17E+03	7.17E+03
Well 25	8.00E+02	3.29E+03	3.48E+03	3.48E+03	3.48E+03	3.48E+03	3.48E+03
Well 26	1.17E+03	1.31E+04	2.04E+04	2.40E+04	2.61E+04	2.75E+04	2.85E+04
Well 27	1.23E+03	1.06E+04	2.35E+04	3.53E+04	4.62E+04	5.64E+04	6.56E+04
Well 28	4.72E+02	3.76E+03	5.02E+03	5.19E+03	5.21E+03	5.22E+03	5.23E+03
Well 29	4.81E+02	4.76E+03	1.18E+04	1.91E+04	2.63E+04	3.34E+04	3.99E+04
Well 30	2.45E+02	1.20E+03	1.30E+03	1.31E+03	1.31E+03	1.31E+03	1.31E+03
Well 31	7.62E+02	3.47E+03	3.78E+03	3.79E+03	3.79E+03	3.79E+03	3.79E+03
Well 32	9.20E+02	1.94E+03	1.96E+03	1.96E+03	1.96E+03	1.96E+03	1.96E+03
Well 33	7.57E+02	4.56E+03	5.47E+03	5.57E+03	5.59E+03	5.59E+03	5.60E+03
Well 34	1.35E+03	1.51E+04	2.29E+04	2.44E+04	2.46E+04	2.47E+04	2.47E+04
Well 35	1.15E+03	5.45E+03	1.05E+04	1.52E+04	1.96E+04	2.37E+04	2.75E+04
Well 36	6.60E+02	3.60E+03	7.12E+03	1.03E+04	1.33E+04	1.61E+04	1.87E+04
Well 37	6.91E+02	3.57E+03	7.18E+03	1.06E+04	1.37E+04	1.68E+04	1.96E+04
Well 38	7.71E+02	1.03E+04	1.59E+04	1.84E+04	1.98E+04	2.06E+04	2.11E+04
Well 39	6.82E+02	1.23E+03	1.23E+03	1.23E+03	1.23E+03	1.23E+03	1.23E+03
Well 40	2.55E+03	4.44E+03	4.45E+03	4.45E+03	4.45E+03	4.45E+03	4.45E+03
Well 41	1.46E+03	2.57E+03	2.57E+03	2.57E+03	2.57E+03	2.57E+03	2.57E+03
Well 42	8.99E+02	2.62E+03	2.69E+03	2.69E+03	2.69E+03	2.69E+03	2.69E+03
Well 43	5.67E+02	1.74E+03	1.91E+03	1.92E+03	1.92E+03	1.92E+03	1.92E+03

A.2. PLED

A.2.1. PLED model parameter

Table A.6 Mean and standard deviation of PLED parameters

Parameters	n		D_i		q_i		D_∞	
Name	μ	σ	μ	σ	μ	σ	μ	σ
Well 1	1.53E-03	1.45E-02	1.97E-03	2.09E-02	1.14E+03	3.10E+01	2.01E-03	3.48E-05
Well 2	8.97E-01	2.62E-02	4.45E-03	7.21E-04	8.58E+02	9.31E+00	4.98E-14	9.34E-14
Well 3	1.40E-01	6.67E-03	6.90E-01	4.32E-02	1.53E+03	7.03E+01	9.00E-14	4.39E-13
Well 4	1.30E+00	5.81E-02	1.33E-03	4.35E-04	6.24E+02	1.38E+01	3.97E-13	4.81E-12
Well 5	8.36E-01	8.06E-02	6.79E-03	3.32E-03	3.20E+02	9.72E+00	1.72E-13	2.74E-12
Well 6	1.57E+00	3.63E-02	1.38E-04	3.16E-05	6.23E+02	5.34E+00	1.42E-12	1.86E-11
Well 7	7.63E-02	1.39E-02	7.65E-01	1.39E-01	8.60E+02	1.17E+02	1.13E-13	1.29E-12
Well 8	2.65E-01	6.03E-02	1.57E-01	8.80E-02	4.52E+02	5.76E+01	1.25E-13	1.50E-12
Well 9	1.26E+00	7.25E-02	1.84E-04	1.39E-04	3.55E+02	2.79E+00	5.14E-14	1.39E-13
Well 10	9.26E-04	4.63E-03	5.36E-04	1.57E-03	8.98E+02	5.95E+00	5.79E-03	5.99E-05
Well 11	8.83E-04	8.21E-03	8.18E-04	2.90E-03	1.02E+03	5.78E+00	5.13E-03	4.16E-05
Well 12	1.43E-03	1.17E-02	8.77E-04	4.81E-03	8.28E+02	5.78E+00	4.51E-03	3.82E-05
Well 13	7.93E-01	3.51E-02	7.79E-03	1.70E-03	1.04E+03	1.53E+01	1.13E-13	1.23E-12
Well 14	5.87E-01	3.48E-02	3.04E-02	6.81E-03	7.48E+02	1.93E+01	9.38E-14	6.78E-13
Well 15	1.96E-01	8.44E-03	5.10E-01	3.61E-02	1.83E+03	8.03E+01	4.99E-14	1.21E-13
Well 16	6.88E-04	2.58E-03	9.56E-04	4.82E-03	5.02E+02	3.27E+00	2.16E-03	2.39E-05
Well 17	1.34E+00	3.37E-02	1.27E-04	3.00E-05	4.67E+02	3.22E+00	5.03E-14	2.09E-13
Well 18	3.01E-01	4.94E-02	1.77E-01	7.23E-02	9.13E+02	1.02E+02	7.69E-14	4.19E-13
Well 19	4.61E-01	6.74E-02	5.43E-02	2.60E-02	8.30E+02	5.04E+01	1.60E-13	3.07E-12
Well 20	1.49E+00	4.08E-02	1.39E-04	3.85E-05	6.48E+02	7.82E+00	1.16E-13	9.66E-13
Well 21	4.22E-01	1.60E-02	2.08E-01	2.29E-02	2.87E+03	1.49E+02	5.26E-14	1.36E-13
Well 22	3.07E-01	6.26E-03	3.43E-01	1.43E-02	3.16E+03	6.73E+01	1.13E-13	6.53E-13
Well 23	8.73E-01	5.32E-02	7.88E-03	2.68E-03	1.27E+03	4.69E+01	1.34E-13	1.18E-12
Well 24	1.41E+00	5.27E-02	1.59E-04	6.03E-05	4.95E+02	4.97E+00	1.05E-13	7.94E-13
Well 25	1.19E-03	1.09E-02	5.63E-03	4.81E-02	7.93E+02	5.32E+01	2.78E-03	4.69E-05
Well 26	7.16E-01	4.35E-02	2.57E-02	7.11E-03	1.22E+03	5.49E+01	1.17E-13	9.20E-13
Well 27	6.07E-01	6.31E-02	3.53E-02	1.38E-02	4.66E+02	2.34E+01	8.90E-14	7.94E-13
Well 28	1.26E-03	8.99E-03	3.75E-03	4.59E-02	2.73E+02	2.06E+01	1.52E-03	5.44E-05
Well 29	1.90E-01	2.89E-02	3.90E-01	1.11E-01	9.34E+02	1.29E+02	1.43E-11	2.45E-10
Well 30	9.88E-02	5.78E-03	8.64E-01	4.99E-02	1.64E+03	8.02E+01	1.49E-03	5.62E-05
Well 31	7.49E-02	1.89E-02	7.75E-01	1.63E-01	1.37E+03	2.14E+02	1.64E-03	1.23E-04
Well 32	1.09E-03	6.61E-03	2.42E-03	2.96E-02	7.65E+02	3.43E+01	4.59E-03	7.46E-05
Well 33	1.34E+00	5.02E-02	2.88E-04	8.73E-05	6.60E+02	5.85E+00	8.78E-14	2.97E-13
Well 34	1.00E-03	6.94E-03	3.67E-03	3.57E-02	1.03E+03	5.01E+01	1.31E-03	4.35E-05
Well 35	2.68E-01	6.27E-02	3.04E-01	1.42E-01	6.30E+02	1.24E+02	3.15E-13	1.75E-11
Well 36	1.91E-01	8.14E-03	4.72E-01	3.37E-02	1.97E+03	8.17E+01	7.63E-14	3.73E-13
Well 37	2.20E-01	5.60E-03	4.52E-01	2.11E-02	2.74E+03	7.60E+01	8.30E-14	3.58E-13
Well 38	8.47E-01	2.33E-02	9.68E-03	1.34E-03	1.08E+03	1.45E+01	4.38E-14	1.63E-13
Well 39	1.01E-03	8.35E-03	1.13E-03	9.32E-03	9.27E+02	1.21E+01	4.82E-03	5.37E-05
Well 40	1.66E+00	2.26E-02	1.18E-04	1.60E-05	1.24E+03	1.71E+01	6.99E-14	2.62E-13
Well 41	1.14E-03	8.84E-03	2.56E-03	2.50E-02	1.14E+03	3.86E+01	4.89E-03	8.28E-05
Well 42	1.36E+00	5.56E-02	4.09E-04	1.36E-04	9.06E+02	1.16E+01	8.83E-14	5.53E-13
Well 43	1.12E-03	9.34E-03	3.33E-03	3.38E-02	3.98E+02	1.88E+01	1.29E-03	4.69E-05

A.2.2. Expected values of PLED daily basis production

Table A.7 Mean of daily production associated with PLED model for several time intervals

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	5.54E+02	2.93E+01	7.55E-01	1.96E-02	5.10E-04	1.33E-05	3.50E-07
Well 2	3.62E+02	2.13E+01	9.28E-01	5.11E-02	3.36E-03	2.57E-04	2.24E-05
Well 3	3.17E+02	2.12E+02	1.73E+02	1.52E+02	1.39E+02	1.28E+02	1.20E+02
Well 4	4.51E+01	3.11E-06	1.63E-13	3.60E-20	6.73E-27	1.02E-33	1.29E-40
Well 5	1.39E+02	1.35E+01	1.51E+00	2.66E-01	6.34E-02	1.87E-02	6.50E-03
Well 6	1.55E+02	2.54E-05	5.47E-16	5.98E-27	3.09E-39	9.19E-53	2.22E-67
Well 7	2.61E+02	2.24E+02	2.08E+02	1.99E+02	1.93E+02	1.88E+02	1.84E+02
Well 8	2.32E+02	1.64E+02	1.35E+02	1.18E+02	1.06E+02	9.76E+01	9.07E+01
Well 9	2.71E+02	4.52E+01	3.39E+00	3.56E-01	7.01E-02	2.16E-02	8.62E-03
Well 10	1.12E+02	2.35E-02	6.21E-07	1.66E-11	4.51E-16	1.24E-20	3.45E-25
Well 11	1.61E+02	8.79E-02	7.65E-06	6.70E-10	5.90E-14	5.22E-18	4.65E-22
Well 12	1.66E+02	2.22E-01	5.98E-05	1.62E-08	4.41E-12	1.21E-15	3.32E-19
Well 13	4.62E+02	5.53E+01	6.75E+00	1.08E+00	2.08E-01	4.58E-02	1.13E-02
Well 14	2.93E+02	6.63E+01	2.00E+01	7.79E+00	3.48E+00	1.71E+00	8.98E-01
Well 15	3.65E+02	1.99E+02	1.45E+02	1.17E+02	1.00E+02	8.79E+01	7.87E+01
Well 16	2.32E+02	9.74E+00	1.89E-01	3.68E-03	7.19E-05	1.40E-06	2.75E-08
Well 17	3.35E+02	2.53E+01	3.62E-01	4.32E-03	8.18E-05	2.42E-06	8.72E-08
Well 18	3.44E+02	1.88E+02	1.32E+02	1.03E+02	8.53E+01	7.28E+01	6.35E+01
Well 19	3.95E+02	1.75E+02	9.89E+01	6.51E+01	4.65E+01	3.49E+01	2.71E+01
Well 20	2.73E+02	5.18E-02	1.04E-07	5.47E-13	2.51E-18	7.43E-24	1.28E-29
Well 21	2.41E+02	2.11E+01	4.02E+00	1.20E+00	4.44E-01	1.90E-01	8.92E-02
Well 22	3.91E+02	1.01E+02	4.47E+01	2.54E+01	1.63E+01	1.12E+01	8.12E+00
Well 23	3.58E+02	7.19E+00	1.36E-01	4.85E-03	2.78E-04	2.31E-05	2.60E-06
Well 24	2.70E+02	1.37E+00	5.10E-04	6.04E-07	1.40E-09	3.86E-12	1.05E-14
Well 25	2.90E+02	4.96E+00	3.15E-02	2.01E-04	1.30E-06	8.42E-09	5.51E-11
Well 26	2.28E+02	5.95E+00	2.34E-01	1.75E-02	1.95E-03	2.89E-04	5.37E-05
Well 27	1.45E+02	2.13E+01	4.75E+00	1.53E+00	6.11E-01	2.80E-01	1.41E-01
Well 28	1.57E+02	1.70E+01	1.07E+00	6.80E-02	4.37E-03	2.84E-04	1.86E-05
Well 29	2.92E+02	1.94E+02	1.56E+02	1.35E+02	1.22E+02	1.12E+02	1.04E+02
Well 30	2.05E+02	1.76E+01	1.02E+00	6.33E-02	4.03E-03	2.63E-04	1.75E-05
Well 31	2.31E+02	1.82E+01	9.12E-01	4.92E-02	2.83E-03	1.72E-04	1.11E-05
Well 32	1.47E+02	1.77E-01	4.19E-05	1.01E-08	2.47E-12	6.18E-16	1.57E-19
Well 33	3.17E+02	1.14E+00	2.26E-04	8.84E-08	8.67E-11	1.08E-13	1.33E-16
Well 34	6.38E+02	9.39E+01	8.68E+00	8.09E-01	7.57E-02	7.14E-03	6.77E-04
Well 35	1.65E+02	8.23E+01	5.55E+01	4.27E+01	3.49E+01	2.97E+01	2.58E+01
Well 36	4.64E+02	2.74E+02	2.07E+02	1.73E+02	1.51E+02	1.35E+02	1.23E+02
Well 37	5.29E+02	2.61E+02	1.77E+02	1.37E+02	1.13E+02	9.63E+01	8.40E+01
Well 38	2.68E+02	4.23E+00	5.55E-02	1.15E-03	3.26E-05	1.19E-06	5.35E-08
Well 39	1.67E+02	1.41E-01	2.17E-05	3.37E-09	5.29E-13	8.37E-17	1.34E-20
Well 40	1.55E+02	2.32E-10	2.37E-30	2.34E-55	2.55E-85	1.18E-119	6.31E-158
Well 41	1.96E+02	1.52E-01	2.08E-05	2.90E-09	4.14E-13	6.04E-17	9.01E-21
Well 42	2.88E+02	4.07E-02	1.46E-07	1.09E-12	8.76E-18	6.02E-23	3.52E-28
Well 43	2.49E+02	3.75E+01	3.58E+00	3.44E-01	3.32E-02	3.24E-03	3.17E-04

A.2.3. Standard deviation values of PLED daily basis production

Table A.8 Standard deviation of daily production regarding PLED model for 7 years

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	4.18E+00	1.71E+00	9.29E-02	3.69E-03	1.30E-04	4.30E-06	1.37E-07
Well 2	2.17E+00	3.08E+00	3.58E-01	3.45E-02	3.47E-03	3.86E-04	4.80E-05
Well 3	2.68E+00	4.57E+00	5.16E+00	5.40E+00	5.52E+00	5.59E+00	5.62E+00
Well 4	4.25E+00	2.24E-05	7.33E-12	1.93E-18	3.78E-25	5.85E-32	7.55E-39
Well 5	4.71E+00	5.81E+00	1.53E+00	4.55E-01	1.65E-01	6.99E-02	3.34E-02
Well 6	4.04E+00	5.89E-05	3.41E-14	5.72E-25	3.07E-37	9.18E-51	2.22E-65
Well 7	2.32E+00	4.62E+00	5.66E+00	6.27E+00	6.70E+00	7.03E+00	7.30E+00
Well 8	2.18E+00	9.57E+00	1.34E+01	1.55E+01	1.68E+01	1.77E+01	1.83E+01
Well 9	2.77E+00	1.30E+01	4.40E+00	1.43E+00	6.09E-01	3.08E-01	1.72E-01
Well 10	1.96E+00	2.51E-03	1.37E-07	5.63E-12	2.09E-16	7.41E-21	2.57E-25
Well 11	1.88E+00	6.36E-03	1.14E-06	1.52E-10	1.81E-14	2.04E-18	2.22E-22
Well 12	1.74E+00	1.47E-02	8.16E-06	3.36E-09	1.24E-12	4.28E-16	1.43E-19
Well 13	4.33E+00	8.80E+00	2.59E+00	6.81E-01	1.88E-01	5.60E-02	1.81E-02
Well 14	3.56E+00	8.23E+00	5.18E+00	3.00E+00	1.76E+00	1.07E+00	6.68E-01
Well 15	2.88E+00	5.18E+00	5.56E+00	5.57E+00	5.48E+00	5.35E+00	5.22E+00
Well 16	1.40E+00	3.88E-01	1.57E-02	4.64E-04	1.22E-05	3.00E-07	7.10E-09
Well 17	3.70E+00	6.08E+00	3.90E-01	1.66E-02	7.88E-04	3.91E-05	1.92E-06
Well 18	4.78E+00	1.61E+01	1.94E+01	2.02E+01	2.02E+01	1.98E+01	1.93E+01
Well 19	5.41E+00	2.28E+01	2.47E+01	2.28E+01	2.03E+01	1.80E+01	1.59E+01
Well 20	6.44E+00	5.75E-02	1.46E-06	2.31E-11	1.79E-16	6.56E-22	1.22E-27
Well 21	3.03E+00	2.05E+00	7.33E-01	3.03E-01	1.41E-01	7.09E-02	3.81E-02
Well 22	4.36E+00	4.21E+00	2.94E+00	2.14E+00	1.62E+00	1.27E+00	1.02E+00
Well 23	5.55E+00	2.85E+00	1.49E-01	1.07E-02	1.19E-03	1.89E-04	3.79E-05
Well 24	4.30E+00	9.59E-01	2.87E-03	1.16E-05	4.75E-08	1.74E-10	5.57E-13
Well 25	3.39E+00	3.99E-01	5.29E-03	5.19E-05	4.54E-07	3.77E-09	3.03E-11
Well 26	5.22E+00	2.05E+00	1.94E-01	2.52E-02	4.46E-03	1.01E-03	2.78E-04
Well 27	5.78E+00	6.84E+00	3.06E+00	1.46E+00	7.65E-01	4.36E-01	2.65E-01
Well 28	2.04E+00	1.56E+00	2.06E-01	2.01E-02	1.77E-03	1.47E-04	1.19E-05
Well 29	2.51E+00	7.78E+00	1.00E+01	1.11E+01	1.17E+01	1.21E+01	1.24E+01
Well 30	1.43E+00	1.36E+00	1.79E-01	1.75E-02	1.55E-03	1.31E-04	1.08E-05
Well 31	3.79E+00	3.21E+00	3.71E-01	3.33E-02	2.86E-03	2.49E-04	2.26E-05
Well 32	2.99E+00	2.26E-02	1.12E-05	4.17E-09	1.42E-12	4.68E-16	1.52E-19
Well 33	3.72E+00	6.41E-01	8.55E-04	1.54E-06	2.85E-09	4.54E-12	6.24E-15
Well 34	6.00E+00	6.74E+00	1.31E+00	1.88E-01	2.39E-02	2.87E-03	3.33E-04
Well 35	4.63E+00	1.17E+01	1.28E+01	1.27E+01	1.22E+01	1.17E+01	1.11E+01
Well 36	2.33E+00	4.88E+00	5.60E+00	5.83E+00	5.90E+00	5.90E+00	5.87E+00
Well 37	2.42E+00	4.31E+00	4.51E+00	4.40E+00	4.23E+00	4.05E+00	3.88E+00
Well 38	3.73E+00	9.36E-01	3.10E-02	1.12E-03	4.93E-05	2.64E-06	1.68E-07
Well 39	2.26E+00	1.28E-02	4.12E-06	9.82E-10	2.10E-13	4.26E-17	8.42E-21
Well 40	8.52E+00	2.11E-09	2.12E-28	2.33E-53	2.55E-83	1.18E-117	6.31E-156
Well 41	4.36E+00	2.13E-02	6.08E-06	1.32E-09	2.61E-13	5.01E-17	9.51E-21
Well 42	5.38E+00	4.33E-02	1.06E-06	2.06E-11	3.03E-16	3.26E-21	2.53E-26
Well 43	2.79E+00	2.96E+00	5.89E-01	8.65E-02	1.14E-02	1.41E-03	1.69E-04

A.2.4. Expected values of PLED cumulative production

Table A.9 Mean of cumulative production corresponding to PLED model

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	2.91E+05	5.53E+05	5.67E+05	5.67E+05	5.67E+05	5.67E+05	5.67E+05
Well 2	2.01E+05	3.72E+05	3.83E+05	3.84E+05	3.84E+05	3.84E+05	3.84E+05
Well 3	1.40E+05	5.09E+05	8.57E+05	1.15E+06	1.42E+06	1.66E+06	1.88E+06
Well 4	9.38E+04	9.83E+04	9.83E+04	9.83E+04	9.83E+04	9.83E+04	9.83E+04
Well 5	7.48E+04	1.49E+05	1.59E+05	1.60E+05	1.60E+05	1.61E+05	1.61E+05
Well 6	1.41E+05	1.63E+05	1.63E+05	1.63E+05	1.63E+05	1.63E+05	1.63E+05
Well 7	1.01E+05	4.52E+05	8.47E+05	1.22E+06	1.58E+06	1.93E+06	2.26E+06
Well 8	9.56E+04	3.76E+05	6.47E+05	8.78E+05	1.08E+06	1.27E+06	1.44E+06
Well 9	1.14E+05	3.09E+05	3.38E+05	3.40E+05	3.40E+05	3.40E+05	3.40E+05
Well 10	1.36E+05	1.55E+05	1.55E+05	1.55E+05	1.55E+05	1.55E+05	1.55E+05
Well 11	1.67E+05	1.98E+05	1.98E+05	1.98E+05	1.98E+05	1.98E+05	1.98E+05
Well 12	1.47E+05	1.84E+05	1.84E+05	1.84E+05	1.84E+05	1.84E+05	1.84E+05
Well 13	2.43E+05	5.08E+05	5.48E+05	5.53E+05	5.54E+05	5.54E+05	5.55E+05
Well 14	1.53E+05	3.60E+05	4.29E+05	4.52E+05	4.61E+05	4.66E+05	4.68E+05
Well 15	1.75E+05	5.54E+05	8.62E+05	1.10E+06	1.30E+06	1.47E+06	1.61E+06
Well 16	1.25E+05	2.28E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05
Well 17	1.46E+05	3.40E+05	3.50E+05	3.50E+05	3.50E+05	3.50E+05	3.50E+05
Well 18	1.57E+05	5.17E+05	8.03E+05	1.01E+06	1.19E+06	1.33E+06	1.45E+06
Well 19	1.81E+05	5.57E+05	7.95E+05	9.41E+05	1.04E+06	1.11E+06	1.17E+06
Well 20	1.70E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05
Well 21	2.12E+05	3.23E+05	3.41E+05	3.45E+05	3.46E+05	3.47E+05	3.47E+05
Well 22	2.49E+05	5.28E+05	6.50E+05	7.11E+05	7.48E+05	7.73E+05	7.90E+05
Well 23	2.47E+05	3.70E+05	3.73E+05	3.73E+05	3.73E+05	3.73E+05	3.73E+05
Well 24	1.40E+05	2.30E+05	2.30E+05	2.30E+05	2.30E+05	2.30E+05	2.30E+05
Well 25	1.79E+05	2.82E+05	2.84E+05	2.84E+05	2.84E+05	2.84E+05	2.84E+05
Well 26	1.83E+05	2.61E+05	2.64E+05	2.64E+05	2.64E+05	2.64E+05	2.64E+05
Well 27	8.45E+04	1.70E+05	1.89E+05	1.94E+05	1.96E+05	1.97E+05	1.97E+05
Well 28	7.52E+04	1.68E+05	1.78E+05	1.79E+05	1.79E+05	1.79E+05	1.79E+05
Well 29	1.27E+05	4.66E+05	7.82E+05	1.05E+06	1.28E+06	1.50E+06	1.69E+06
Well 30	1.15E+05	2.23E+05	2.34E+05	2.34E+05	2.34E+05	2.34E+05	2.34E+05
Well 31	1.23E+05	2.43E+05	2.54E+05	2.54E+05	2.54E+05	2.54E+05	2.54E+05
Well 32	1.34E+05	1.66E+05	1.66E+05	1.66E+05	1.66E+05	1.66E+05	1.66E+05
Well 33	1.77E+05	2.74E+05	2.75E+05	2.75E+05	2.75E+05	2.75E+05	2.75E+05
Well 34	2.92E+05	7.09E+05	7.73E+05	7.79E+05	7.80E+05	7.80E+05	7.80E+05
Well 35	7.97E+04	2.44E+05	3.66E+05	4.55E+05	5.26E+05	5.86E+05	6.35E+05
Well 36	2.14E+05	7.18E+05	1.15E+06	1.50E+06	1.79E+06	2.05E+06	2.28E+06
Well 37	2.64E+05	7.84E+05	1.17E+06	1.46E+06	1.68E+06	1.87E+06	2.03E+06
Well 38	1.95E+05	2.82E+05	2.84E+05	2.84E+05	2.84E+05	2.84E+05	2.84E+05
Well 39	1.58E+05	1.92E+05	1.92E+05	1.92E+05	1.92E+05	1.92E+05	1.92E+05
Well 40	2.43E+05	2.57E+05	2.57E+05	2.57E+05	2.57E+05	2.57E+05	2.57E+05
Well 41	1.92E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05	2.32E+05
Well 42	2.12E+05	2.70E+05	2.70E+05	2.70E+05	2.70E+05	2.70E+05	2.70E+05
Well 43	1.14E+05	2.78E+05	3.04E+05	3.07E+05	3.07E+05	3.07E+05	3.07E+05

A.2.5. Standard deviation values of PLED cumulative production

Table A.10 Standard deviation of cumulative production respected with PLED model

Name	Year 1	Year 5	Year 10	Year 15	Year 20	Year 25	Year 30
Well 1	9.98E+02	5.68E+03	6.63E+03	6.68E+03	6.68E+03	6.68E+03	6.68E+03
Well 2	4.00E+02	6.74E+03	9.09E+03	9.33E+03	9.35E+03	9.35E+03	9.35E+03
Well 3	6.65E+02	6.21E+03	1.53E+04	2.52E+04	3.55E+04	4.59E+04	5.59E+04
Well 4	1.11E+03	1.74E+03	1.74E+03	1.74E+03	1.74E+03	1.74E+03	1.74E+03
Well 5	8.29E+02	1.12E+04	1.68E+04	1.81E+04	1.85E+04	1.86E+04	1.87E+04
Well 6	7.69E+02	1.82E+03	1.82E+03	1.82E+03	1.82E+03	1.82E+03	1.82E+03
Well 7	5.73E+02	5.88E+03	1.54E+04	2.64E+04	3.84E+04	5.11E+04	6.38E+04
Well 8	4.04E+02	9.76E+03	3.14E+04	5.83E+04	8.82E+04	1.20E+05	1.52E+05
Well 9	4.60E+02	1.56E+04	2.95E+04	3.29E+04	3.37E+04	3.39E+04	3.40E+04
Well 10	6.93E+02	1.14E+03	1.14E+03	1.14E+03	1.14E+03	1.14E+03	1.14E+03
Well 11	6.16E+02	1.10E+03	1.11E+03	1.11E+03	1.11E+03	1.11E+03	1.11E+03
Well 12	5.24E+02	1.07E+03	1.07E+03	1.07E+03	1.07E+03	1.07E+03	1.07E+03
Well 13	7.56E+02	1.42E+04	2.34E+04	2.59E+04	2.65E+04	2.67E+04	2.67E+04
Well 14	6.54E+02	1.10E+04	2.29E+04	2.99E+04	3.39E+04	3.62E+04	3.76E+04
Well 15	7.31E+02	6.83E+03	1.67E+04	2.68E+04	3.68E+04	4.67E+04	5.59E+04
Well 16	3.49E+02	1.68E+03	1.89E+03	1.90E+03	1.90E+03	1.90E+03	1.90E+03
Well 17	6.63E+02	1.15E+04	1.52E+04	1.54E+04	1.54E+04	1.54E+04	1.54E+04
Well 18	1.08E+03	1.79E+04	5.12E+04	8.77E+04	1.25E+05	1.61E+05	1.95E+05
Well 19	1.02E+03	2.48E+04	6.95E+04	1.13E+05	1.52E+05	1.86E+05	2.15E+05
Well 20	1.28E+03	4.99E+03	5.00E+03	5.00E+03	5.00E+03	5.00E+03	5.00E+03
Well 21	9.19E+02	5.12E+03	7.44E+03	8.32E+03	8.70E+03	8.89E+03	8.98E+03
Well 22	1.20E+03	7.05E+03	1.31E+04	1.74E+04	2.06E+04	2.31E+04	2.50E+04
Well 23	1.44E+03	1.12E+04	1.27E+04	1.28E+04	1.28E+04	1.28E+04	1.28E+04
Well 24	7.69E+02	6.87E+03	7.08E+03	7.08E+03	7.08E+03	7.08E+03	7.08E+03
Well 25	8.28E+02	3.10E+03	3.25E+03	3.25E+03	3.25E+03	3.25E+03	3.25E+03
Well 26	1.15E+03	7.79E+03	9.14E+03	9.27E+03	9.29E+03	9.30E+03	9.30E+03
Well 27	1.17E+03	1.27E+04	2.17E+04	2.58E+04	2.78E+04	2.89E+04	2.95E+04
Well 28	4.42E+02	3.86E+03	5.14E+03	5.28E+03	5.29E+03	5.30E+03	5.30E+03
Well 29	4.99E+02	8.45E+03	2.48E+04	4.40E+04	6.48E+04	8.64E+04	1.08E+05
Well 30	3.11E+02	2.93E+03	3.98E+03	4.10E+03	4.11E+03	4.11E+03	4.11E+03
Well 31	7.44E+02	7.88E+03	1.03E+04	1.06E+04	1.06E+04	1.06E+04	1.06E+04
Well 32	9.15E+02	1.85E+03	1.85E+03	1.85E+03	1.85E+03	1.85E+03	1.85E+03
Well 33	6.45E+02	6.97E+03	7.14E+03	7.14E+03	7.14E+03	7.14E+03	7.14E+03
Well 34	1.29E+03	1.32E+04	1.95E+04	2.05E+04	2.06E+04	2.06E+04	2.06E+04
Well 35	1.10E+03	1.45E+04	3.81E+04	6.22E+04	8.57E+04	1.08E+05	1.29E+05
Well 36	6.09E+02	6.04E+03	1.59E+04	2.65E+04	3.75E+04	4.86E+04	5.92E+04
Well 37	6.47E+02	5.66E+03	1.38E+04	2.19E+04	2.98E+04	3.74E+04	4.43E+04
Well 38	7.15E+02	4.74E+03	5.20E+03	5.22E+03	5.22E+03	5.22E+03	5.22E+03
Well 39	7.01E+02	1.23E+03	1.24E+03	1.24E+03	1.24E+03	1.24E+03	1.24E+03
Well 40	2.54E+03	3.28E+03	3.28E+03	3.28E+03	3.28E+03	3.28E+03	3.28E+03
Well 41	1.44E+03	2.48E+03	2.48E+03	2.48E+03	2.48E+03	2.48E+03	2.48E+03
Well 42	9.13E+02	4.89E+03	4.90E+03	4.90E+03	4.90E+03	4.90E+03	4.90E+03
Well 43	5.51E+02	5.86E+03	8.59E+03	9.05E+03	9.12E+03	9.13E+03	9.13E+03

APPENDIX B: TASK 2, CROSS-VALIDATION RESULTS

B.1. MHD

B.1.1. Model parameters

Variable b results are presented in the Cross-validation results.

Table B.1 MHD parameter D_i

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.47E-03	1.39E-04	4.04E-04	4.32E-05	3.24E-03	3.66E-04	2.96E-04	3.34E-05
$\boldsymbol{OK_{Sp}}$	1.35E-03	1.29E-04	3.49E-04	4.23E-05	3.15E-03	3.55E-04	2.87E-04	3.24E-05
$\boldsymbol{OK_{ex}}$	1.35E-03	1.27E-04	3.49E-04	4.14E-05	2.96E-03	3.35E-04	2.70E-04	3.06E-05
$\boldsymbol{OK_{Ga}}$	5.30E-03	3.03E-04	7.94E-04	6.70E-05	3.78E-03	5.78E-04	3.54E-04	5.27E-05
$\boldsymbol{OK_{St}}$	1.37E-03	1.30E-04	3.48E-04	4.22E-05	3.17E-03	3.58E-04	2.89E-04	3.27E-05

Table B.2 MHD parameter q_i

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	8.04E+01	2.73E+00	6.75E+01	1.20E+00	9.94E+01	6.28E+00	1.22E+01	5.84E-01
$\boldsymbol{OK_{Sp}}$	7.78E+01	2.52E+00	6.31E+01	1.23E+00	1.03E+02	6.17E+00	1.21E+01	5.76E-01
$\boldsymbol{OK_{ex}}$	7.66E+01	2.49E+00	6.32E+01	1.19E+00	1.04E+02	5.83E+00	1.22E+01	5.45E-01
$\boldsymbol{OK_{Ga}}$	4.55E+02	7.40E+00	2.71E+02	7.40E+00	1.14E+03	1.09E+02	1.08E+02	9.85E+00
$\boldsymbol{OK_{St}}$	8.77E+01	2.68E+00	6.58E+01	1.29E+00	1.03E+02	6.22E+00	1.23E+01	5.82E-01

B.1.2. Daily basis production

Table B.3 Daily basis production MHD 1 year

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	2.52E+01	2.81E-01	3.36E+01	3.75E-01	3.37E+01	4.62E-01	5.13E+00	6.21E-02
$\boldsymbol{OK_{Sp}}$	2.42E+01	2.91E-01	2.77E+01	3.57E-01	3.05E+01	4.63E-01	4.38E+00	6.06E-02
$\boldsymbol{OK_{ex}}$	2.40E+01	2.84E-01	2.75E+01	3.56E-01	2.98E+01	4.44E-01	4.33E+00	5.93E-02
$\boldsymbol{OK_{Ga}}$	2.06E+02	4.61E+01	2.97E+01	2.20E+00	2.73E+02	1.34E+03	2.49E+01	1.21E+02
$\boldsymbol{OK_{St}}$	2.43E+01	4.09E-01	2.76E+01	3.74E-01	3.18E+01	4.81E-01	4.44E+00	6.32E-02

Table B.4 Daily basis production MHD 5 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.16E+01	8.84E-01	1.48E+01	1.07E+00	2.70E+01	1.44E+00	3.04E+00	1.85E-01
$\boldsymbol{OK_{Sp}}$	1.13E+01	8.30E-01	1.13E+01	1.01E+00	2.50E+01	1.39E+00	2.65E+00	1.77E-01
$\boldsymbol{OK_{ex}}$	1.11E+01	8.24E-01	1.08E+01	9.95E-01	2.45E+01	1.38E+00	2.58E+00	1.74E-01
$\boldsymbol{OK_{Ga}}$	4.20E+02	1.38E+00	1.10E+01	1.57E+00	2.92E+02	5.13E+00	2.63E+01	5.01E-01
$\boldsymbol{OK_{St}}$	1.68E+01	8.39E-01	1.10E+01	1.02E+00	2.57E+01	1.50E+00	2.68E+00	1.84E-01

Table B.5 Daily basis production MHD 10 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	7.01E+00	6.64E-01	8.27E+00	7.48E-01	1.60E+01	9.70E-01	1.77E+00	1.27E-01
$\boldsymbol{OK_{Sp}}$	6.98E+00	6.28E-01	6.02E+00	7.25E-01	1.47E+01	9.39E-01	1.52E+00	1.23E-01
$\boldsymbol{OK_{ex}}$	6.86E+00	6.22E-01	5.88E+00	7.11E-01	1.44E+01	9.27E-01	1.49E+00	1.21E-01
$\boldsymbol{OK_{Ga}}$	2.43E+02	9.12E-01	5.77E+00	9.37E-01	1.69E+02	2.68E+00	1.52E+01	2.68E-01
$\boldsymbol{OK_{St}}$	1.06E+01	6.34E-01	5.80E+00	7.22E-01	1.51E+01	1.04E+00	1.54E+00	1.29E-01

Table B.6 Daily basis production MHD 15 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	4.19E+00	4.11E-01	4.82E+00	4.53E-01	9.42E+00	5.66E-01	1.04E+00	7.56E-02
$\boldsymbol{OK_{Sp}}$	4.18E+00	3.89E-01	3.53E+00	4.41E-01	8.63E+00	5.48E-01	8.90E-01	7.34E-02
$\boldsymbol{OK_{ex}}$	4.11E+00	3.86E-01	3.45E+00	4.33E-01	8.49E+00	5.41E-01	8.75E-01	7.22E-02
$\boldsymbol{OK_{Ga}}$	1.43E+02	5.94E-01	3.38E+00	4.96E-01	1.01E+02	1.57E+00	9.08E+00	1.54E-01
$\boldsymbol{OK_{St}}$	6.28E+00	3.94E-01	3.39E+00	4.37E-01	8.90E+00	6.25E-01	9.04E-01	7.78E-02

Table B.7 Daily basis production MHD 20 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	2.48E+00	2.43E-01	2.84E+00	2.66E-01	5.56E+00	3.32E-01	6.10E-01	4.43E-02
$\boldsymbol{OK_{Sp}}$	2.48E+00	2.30E-01	2.08E+00	2.59E-01	5.09E+00	3.22E-01	5.26E-01	4.31E-02
$\boldsymbol{OK_{ex}}$	2.44E+00	2.28E-01	2.05E+00	2.54E-01	5.01E+00	3.18E-01	5.17E-01	4.24E-02
$\boldsymbol{OK_{Ga}}$	8.20E+01	3.60E-01	2.01E+00	2.75E-01	5.97E+01	9.23E-01	5.38E+00	8.98E-02
$\boldsymbol{OK_{St}}$	3.69E+00	2.33E-01	2.01E+00	2.55E-01	5.26E+00	3.71E-01	5.34E-01	4.59E-02

Table B.8 Daily basis production MHD 25 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.47E+00	1.43E-01	1.67E+00	1.56E-01	3.28E+00	1.96E-01	3.60E-01	2.60E-02
$\boldsymbol{OK_{Sp}}$	1.47E+00	1.36E-01	1.23E+00	1.52E-01	3.00E+00	1.90E-01	3.10E-01	2.53E-02
$\boldsymbol{OK_{ex}}$	1.44E+00	1.34E-01	1.21E+00	1.49E-01	2.96E+00	1.87E-01	3.05E-01	2.49E-02
$\boldsymbol{OK_{Ga}}$	4.72E+01	2.16E-01	1.19E+00	1.59E-01	3.53E+01	5.45E-01	3.18E+00	5.28E-02
$\boldsymbol{OK_{St}}$	2.17E+00	1.37E-01	1.19E+00	1.49E-01	3.11E+00	2.20E-01	3.16E-01	2.70E-02

Table B.9 Daily basis production MHD 30 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	8.66E-01	8.43E-02	9.87E-01	9.13E-02	1.94E+00	1.16E-01	2.13E-01	1.53E-02
$\boldsymbol{OK_{Sp}}$	8.66E-01	8.02E-02	7.30E-01	8.93E-02	1.77E+00	1.12E-01	1.83E-01	1.49E-02
$\boldsymbol{OK_{ex}}$	8.51E-01	7.91E-02	7.17E-01	8.75E-02	1.75E+00	1.10E-01	1.80E-01	1.47E-02
$\boldsymbol{OK_{Ga}}$	2.74E+01	1.29E-01	7.06E-01	9.26E-02	2.09E+01	3.22E-01	1.88E+00	3.11E-02
$\boldsymbol{OK_{St}}$	1.28E+00	8.08E-02	7.06E-01	8.77E-02	1.83E+00	1.30E-01	1.87E-01	1.59E-02

B.1.3. Cumulative production

Table B.10 Cumulative production MHD 1 year

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.20E+04	7.56E+01	1.56E+04	6.91E+01	9.57E+03	1.67E+02	2.10E+03	1.72E+01
$\boldsymbol{OK_{Sp}}$	1.20E+04	8.34E+01	1.44E+04	6.77E+01	1.05E+04	1.72E+02	2.01E+03	1.76E+01
$\boldsymbol{OK_{ex}}$	1.17E+04	7.96E+01	1.38E+04	6.67E+01	1.07E+04	1.67E+02	1.96E+03	1.72E+01
$\boldsymbol{OK_{Ga}}$	2.84E+04	6.21E+04	3.25E+04	4.63E+02	1.10E+05	2.42E+06	1.07E+04	2.18E+05
$\boldsymbol{OK_{St}}$	1.24E+04	1.40E+02	1.38E+04	7.21E+01	1.28E+04	2.64E+02	2.05E+03	2.54E+01

Table B.11 Cumulative production MHD 5 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	3.09E+04	1.05E+03	4.19E+04	1.41E+03	5.03E+04	1.40E+03	6.87E+03	2.14E+02
$\boldsymbol{OK_{Sp}}$	2.98E+04	9.94E+02	3.39E+04	1.31E+03	4.91E+04	1.37E+03	6.08E+03	2.02E+02
$\boldsymbol{OK_{ex}}$	2.97E+04	9.92E+02	3.26E+04	1.30E+03	4.89E+04	1.34E+03	5.96E+03	2.01E+02
$\boldsymbol{OK_{Ga}}$	6.38E+05	2.63E+03	3.59E+04	2.06E+03	6.35E+05	1.02E+04	5.73E+04	9.50E+02
$\boldsymbol{OK_{St}}$	3.57E+04	1.00E+03	3.32E+04	1.33E+03	5.00E+04	1.50E+03	6.08E+03	2.12E+02

Table B.12 Cumulative production MHD 10 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	4.39E+04	2.49E+03	5.84E+04	2.99E+03	8.94E+04	4.14E+03	1.08E+04	5.24E+02
$\boldsymbol{OK_{Sp}}$	4.28E+04	2.33E+03	4.50E+04	2.80E+03	8.49E+04	4.03E+03	9.44E+03	5.01E+02
$\boldsymbol{OK_{ex}}$	4.22E+04	2.32E+03	4.30E+04	2.77E+03	8.39E+04	3.99E+03	9.22E+03	4.95E+02
$\boldsymbol{OK_{Ga}}$	2.81E+06	4.55E+03	4.79E+04	4.14E+03	4.23E+06	1.98E+04	3.81E+05	1.86E+03
$\boldsymbol{OK_{St}}$	6.09E+04	2.36E+03	4.37E+04	2.83E+03	8.67E+04	4.47E+03	9.48E+03	5.32E+02

Table B.13 Cumulative production MHD 15 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	5.48E+04	3.75E+03	7.15E+04	4.20E+03	1.20E+05	6.85E+03	1.39E+04	8.05E+02
$\boldsymbol{OK_{Sp}}$	5.37E+04	3.49E+03	5.27E+04	3.96E+03	1.13E+05	6.66E+03	1.20E+04	7.73E+02
$\boldsymbol{OK_{ex}}$	5.29E+04	3.48E+03	5.12E+04	3.91E+03	1.11E+05	6.58E+03	1.18E+04	7.63E+02
$\boldsymbol{OK_{Ga}}$	3.63E+06	6.33E+03	5.56E+04	5.53E+03	5.62E+06	2.89E+04	5.06E+05	2.69E+03
$\boldsymbol{OK_{St}}$	8.13E+04	3.53E+03	5.26E+04	3.99E+03	1.15E+05	7.50E+03	1.22E+04	8.35E+02

Table B.14 Cumulative production MHD 20 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	6.45E+04	4.86E+03	8.32E+04	5.20E+03	1.46E+05	9.43E+03	1.66E+04	1.06E+03
$\boldsymbol{OK_{Sp}}$	6.36E+04	4.51E+03	6.15E+04	4.94E+03	1.36E+05	9.17E+03	1.44E+04	1.03E+03
$\boldsymbol{OK_{ex}}$	6.26E+04	4.53E+03	5.89E+04	4.86E+03	1.34E+05	9.06E+03	1.41E+04	1.01E+03
$\boldsymbol{OK_{Ga}}$	3.96E+06	8.23E+03	6.31E+04	6.62E+03	6.71E+06	3.77E+04	6.04E+05	3.49E+03
$\boldsymbol{OK_{St}}$	9.88E+04	4.58E+03	5.92E+04	4.96E+03	1.40E+05	1.05E+04	1.46E+04	1.12E+03

Table B.15 Cumulative production MHD 25 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	7.34E+04	5.87E+03	9.39E+04	6.08E+03	1.69E+05	1.19E+04	1.91E+04	1.31E+03
$\boldsymbol{OK_{Sp}}$	7.27E+04	5.45E+03	6.89E+04	5.81E+03	1.57E+05	1.16E+04	1.65E+04	1.26E+03
$\boldsymbol{OK_{ex}}$	7.14E+04	5.49E+03	6.60E+04	5.71E+03	1.55E+05	1.15E+04	1.61E+04	1.25E+03
$\boldsymbol{OK_{Ga}}$	4.12E+06	1.02E+04	7.01E+04	7.56E+03	7.67E+06	4.67E+04	6.91E+05	4.31E+03
$\boldsymbol{OK_{St}}$	1.14E+05	5.54E+03	6.69E+04	5.82E+03	1.62E+05	1.33E+04	1.67E+04	1.40E+03

Table B.16 Cumulative production MHD 30 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	8.13E+04	6.78E+03	1.04E+05	6.86E+03	1.89E+05	1.42E+04	2.12E+04	1.53E+03
$\boldsymbol{OK_{Sp}}$	8.07E+04	6.29E+03	7.56E+04	6.57E+03	1.76E+05	1.38E+04	1.84E+04	1.48E+03
$\boldsymbol{OK_{ex}}$	7.92E+04	6.37E+03	7.25E+04	6.45E+03	1.73E+05	1.36E+04	1.79E+04	1.46E+03
$\boldsymbol{OK_{Ga}}$	4.20E+06	1.20E+04	7.64E+04	8.38E+03	8.50E+06	5.46E+04	7.66E+05	5.02E+03
$\boldsymbol{OK_{St}}$	1.28E+05	6.43E+03	7.27E+04	6.58E+03	1.81E+05	1.60E+04	1.86E+04	1.65E+03

B.2. PLED

B.2.1. Model parameters

Table B.17 PLED parameter n

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.18E-01	4.25E-03	1.39E-01	5.42E-03	2.25E-01	9.39E-03	2.65E-02	1.08E-03
$\boldsymbol{OK_{Sp}}$	1.08E-01	4.04E-03	1.34E-01	5.10E-03	2.18E-01	9.31E-03	2.57E-02	1.05E-03
$\boldsymbol{OK_{ex}}$	1.07E-01	4.03E-03	1.31E-01	5.02E-03	2.01E-01	9.12E-03	2.43E-02	1.03E-03
$\boldsymbol{OK_{Ga}}$	1.13E+00	5.02E-02	1.34E-01	5.62E-03	7.93E-01	5.87E-02	7.33E-02	5.33E-03
$\boldsymbol{OK_{St}}$	1.33E-01	4.86E-03	1.34E-01	5.08E-03	2.34E-01	9.47E-03	2.68E-02	1.06E-03

Table B.18 PLED parameter D_i

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	4.45E-02	7.54E-03	6.92E-02	1.03E-02	2.24E-02	7.30E-03	8.75E-03	1.43E-03
$\boldsymbol{OK_{Sp}}$	4.60E-02	7.40E-03	5.97E-02	9.74E-03	1.97E-02	6.70E-03	7.56E-03	1.34E-03
$\boldsymbol{OK_{ex}}$	4.48E-02	7.20E-03	5.86E-02	9.63E-03	1.98E-02	6.62E-03	7.43E-03	1.33E-03
$\boldsymbol{OK_{Ga}}$	7.36E-01	3.63E-02	6.25E-02	8.29E-03	4.06E+00	2.18E+00	3.66E-01	1.96E-01
$\boldsymbol{OK_{St}}$	4.63E-02	7.36E-03	6.09E-02	9.42E-03	2.07E-02	7.98E-03	7.73E-03	1.36E-03

Table B.19 PLED parameter q_i

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.22E+02	8.33E+00	1.13E+02	1.05E+01	7.23E+01	1.27E+01	1.54E+01	1.72E+00
$\boldsymbol{OK_{Sp}}$	1.19E+02	8.69E+00	1.00E+02	1.08E+01	7.74E+01	1.16E+01	1.42E+01	1.69E+00
$\boldsymbol{OK_{ex}}$	1.18E+02	8.55E+00	9.02E+01	9.74E+00	9.04E+01	1.15E+01	1.38E+01	1.58E+00
$\boldsymbol{OK_{Ga}}$	1.03E+03	1.61E+03	2.31E+02	8.43E+00	3.21E+03	1.55E+03	2.91E+02	1.40E+02
$\boldsymbol{OK_{St}}$	1.44E+02	9.33E+00	9.88E+01	9.74E+00	7.65E+01	1.28E+01	1.40E+01	1.66E+00

Table B.20 PLED parameter D_∞

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	3.43E-04	5.29E-06	3.33E-04	6.90E-06	8.66E-04	1.11E-05	8.81E-05	1.31E-06
$\boldsymbol{OK_{Sp}}$	3.20E-04	4.71E-06	3.32E-04	7.00E-06	8.21E-04	1.03E-05	8.45E-05	1.27E-06
$\boldsymbol{OK_{ex}}$	3.19E-04	4.70E-06	3.20E-04	6.81E-06	7.95E-04	1.01E-05	8.18E-05	1.24E-06
$\boldsymbol{OK_{Ga}}$	5.43E-03	1.58E-05	8.77E-03	1.35E-05	6.87E-03	5.60E-05	1.24E-03	5.31E-06
$\boldsymbol{OK_{St}}$	3.68E-04	5.97E-06	5.71E-04	1.01E-05	8.23E-04	1.30E-05	1.02E-04	1.71E-06

B.2.2. Daily basis production

Table B.21 Daily basis production PLED 1 year

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	2.52E+01	2.76E-01	3.37E+01	3.23E-01	3.53E+01	5.22E-01	5.23E+00	6.16E-02
$\boldsymbol{OK_{Sp}}$	2.40E+01	2.87E-01	2.75E+01	3.16E-01	3.20E+01	4.09E-01	4.45E+00	5.35E-02
$\boldsymbol{OK_{ex}}$	2.39E+01	2.77E-01	2.73E+01	3.10E-01	3.15E+01	4.17E-01	4.40E+00	5.36E-02
$\boldsymbol{OK_{Ga}}$	2.23E+02	2.09E+01	2.95E+01	2.39E+00	2.42E+02	6.07E+00	2.21E+01	6.21E-01
$\boldsymbol{OK_{St}}$	2.42E+01	4.20E-01	2.74E+01	3.35E-01	3.28E+01	4.76E-01	4.48E+00	5.95E-02

Table B.22 Daily basis production PLED 5 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.32E+01	6.18E-01	1.97E+01	6.96E-01	2.55E+01	2.93E+00	3.34E+00	2.77E-01
$\boldsymbol{OK_{Sp}}$	1.36E+01	6.18E-01	1.47E+01	7.34E-01	2.37E+01	2.79E+00	2.80E+00	2.67E-01
$\boldsymbol{OK_{ex}}$	1.32E+01	6.25E-01	1.41E+01	6.71E-01	2.32E+01	2.68E+00	2.72E+00	2.55E-01
$\boldsymbol{OK_{Ga}}$	2.45E+02	1.18E+01	1.62E+01	8.65E-01	3.32E+02	1.93E+01	3.00E+01	1.74E+00
$\boldsymbol{OK_{St}}$	1.93E+01	6.55E-01	1.47E+01	7.49E-01	2.47E+01	3.14E+00	2.86E+00	2.97E-01

Table B.23 Daily basis production PLED 10 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.12E+01	7.04E-01	1.73E+01	8.74E-01	1.57E+01	2.98E+00	2.56E+00	2.89E-01
$\boldsymbol{OK_{Sp}}$	1.10E+01	7.20E-01	1.33E+01	7.91E-01	1.44E+01	2.80E+00	2.09E+00	2.70E-01
$\boldsymbol{OK_{ex}}$	1.12E+01	7.11E-01	1.28E+01	7.90E-01	1.44E+01	2.79E+00	2.04E+00	2.69E-01
$\boldsymbol{OK_{Ga}}$	1.53E+01	1.95E+01	1.75E+01	9.09E-01	4.82E+05	2.44E+01	4.34E+04	2.20E+00
$\boldsymbol{OK_{St}}$	1.14E+01	8.80E-01	1.34E+01	8.03E-01	1.78E+01	3.41E+00	2.30E+00	3.22E-01

Table B.24 Daily basis production PLED 15 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	9.98E+00	7.61E-01	1.55E+01	1.01E+00	1.14E+01	2.84E+00	2.17E+00	2.84E-01
$\boldsymbol{OK_{Sp}}$	9.43E+00	8.04E-01	1.21E+01	8.86E-01	1.05E+01	2.66E+00	1.76E+00	2.63E-01
$\boldsymbol{OK_{ex}}$	9.73E+00	7.84E-01	1.17E+01	8.83E-01	1.04E+01	2.66E+00	1.72E+00	2.63E-01
$\boldsymbol{OK_{Ga}}$	1.07E+01	9.42E+00	1.75E+01	1.07E+00	4.53E+05	3.00E+01	4.08E+04	2.70E+00
$\boldsymbol{OK_{St}}$	9.67E+00	9.10E-01	1.23E+01	8.96E-01	1.30E+01	3.31E+00	1.91E+00	3.18E-01

Table B.25 Daily basis production PLED 20 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	9.15E+00	7.84E-01	1.42E+01	1.08E+00	9.00E+00	2.65E+00	1.93E+00	2.73E-01
$\boldsymbol{OK_{Sp}}$	8.45E+00	8.28E-01	1.12E+01	9.37E-01	8.19E+00	2.43E+00	1.56E+00	2.47E-01
$\boldsymbol{OK_{ex}}$	8.77E+00	8.13E-01	1.09E+01	9.31E-01	8.21E+00	2.47E+00	1.53E+00	2.50E-01
$\boldsymbol{OK_{Ga}}$	9.09E+00	6.52E+00	1.73E+01	1.18E+00	4.08E+05	9.54E+04	3.68E+04	8.59E+03
$\boldsymbol{OK_{St}}$	8.71E+00	8.88E-01	1.10E+01	9.48E-01	1.02E+01	3.09E+00	1.64E+00	3.02E-01

Table B.26 Daily basis production PLED 25 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	8.55E+00	7.99E-01	1.33E+01	1.11E+00	7.43E+00	2.47E+00	1.76E+00	2.61E-01
$\boldsymbol{OK_{Sp}}$	8.11E+00	8.48E-01	1.05E+01	9.66E-01	6.73E+00	2.29E+00	1.42E+00	2.38E-01
$\boldsymbol{OK_{ex}}$	8.10E+00	8.27E-01	1.02E+01	9.58E-01	6.75E+00	2.30E+00	1.40E+00	2.38E-01
$\boldsymbol{OK_{Ga}}$	8.24E+00	4.43E+00	1.69E+01	1.27E+00	3.65E+05	9.62E+04	3.29E+04	8.67E+03
$\boldsymbol{OK_{St}}$	7.96E+00	8.70E-01	1.08E+01	9.78E-01	8.42E+00	2.87E+00	1.53E+00	2.85E-01

Table B.27 Daily basis production PLED 30 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	8.08E+00	8.10E-01	1.25E+01	1.13E+00	6.33E+00	2.31E+00	1.64E+00	2.50E-01
$\boldsymbol{OK_{Sp}}$	7.36E+00	8.54E-01	9.90E+00	9.83E-01	5.71E+00	2.13E+00	1.32E+00	2.27E-01
$\boldsymbol{OK_{ex}}$	7.59E+00	8.31E-01	9.75E+00	9.73E-01	5.73E+00	2.14E+00	1.31E+00	2.27E-01
$\boldsymbol{OK_{Ga}}$	7.66E+00	3.08E+00	1.65E+01	1.35E+00	3.27E+05	9.29E+04	2.94E+04	8.36E+03
$\boldsymbol{OK_{St}}$	7.39E+00	8.57E-01	1.03E+01	9.97E-01	7.14E+00	2.67E+00	1.42E+00	2.70E-01

B.2.3. Cumulative production

Table B.28 Cumulative production PLED 1 year

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.20E+04	7.64E+01	1.56E+04	6.73E+01	8.89E+03	1.65E+02	2.08E+03	1.70E+01
$\boldsymbol{OK_{Sp}}$	1.19E+04	8.47E+01	1.42E+04	6.75E+01	9.97E+03	1.70E+02	1.97E+03	1.74E+01
$\boldsymbol{OK_{ex}}$	1.17E+04	8.04E+01	1.38E+04	6.59E+01	1.03E+04	1.69E+02	1.93E+03	1.72E+01
$\boldsymbol{OK_{Ga}}$	2.83E+04	6.02E+04	3.17E+04	3.86E+02	3.99E+06	3.26E+06	3.59E+05	2.94E+05
$\boldsymbol{OK_{St}}$	1.23E+04	1.39E+02	1.37E+04	7.17E+01	1.24E+04	2.54E+02	2.02E+03	2.45E+01

Table B.29 Cumulative production PLED 5 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	3.09E+04	7.35E+02	4.31E+04	9.64E+02	5.07E+04	2.89E+03	7.00E+03	2.86E+02
$\boldsymbol{OK_{Sp}}$	3.04E+04	7.35E+02	3.12E+04	9.14E+02	4.89E+04	2.76E+03	5.84E+03	2.73E+02
$\boldsymbol{OK_{ex}}$	3.01E+04	7.44E+02	3.15E+04	9.06E+02	4.86E+04	2.67E+03	5.84E+03	2.65E+02
$\boldsymbol{OK_{Ga}}$	9.04E+05	1.60E+04	3.64E+04	2.34E+03	6.55E+05	1.76E+04	5.92E+04	1.61E+03
$\boldsymbol{OK_{St}}$	4.03E+04	8.12E+02	3.22E+04	9.67E+02	5.11E+04	2.98E+03	6.08E+03	2.94E+02

Table B.30 Cumulative production PLED 10 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	4.74E+04	1.75E+03	7.01E+04	2.07E+03	8.53E+04	8.21E+03	1.15E+04	7.82E+02
$\boldsymbol{OK_{Sp}}$	4.77E+04	1.75E+03	5.14E+04	2.08E+03	7.60E+04	7.81E+03	9.32E+03	7.49E+02
$\boldsymbol{OK_{ex}}$	4.68E+04	1.77E+03	4.86E+04	1.98E+03	8.07E+04	7.68E+03	9.41E+03	7.34E+02
$\boldsymbol{OK_{Ga}}$	1.41E+06	9.05E+04	5.06E+04	2.36E+03	4.05E+06	6.78E+04	3.65E+05	6.11E+03
$\boldsymbol{OK_{St}}$	7.11E+04	2.13E+03	5.04E+04	2.11E+03	9.31E+04	8.87E+03	1.04E+04	8.40E+02

Table B.31 Cumulative production PLED 15 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	6.38E+04	2.93E+03	9.68E+04	3.59E+03	1.09E+05	1.34E+04	1.54E+04	1.28E+03
$\boldsymbol{OK_{Sp}}$	6.51E+04	2.99E+03	7.12E+04	3.43E+03	1.03E+05	1.27E+04	1.28E+04	1.22E+03
$\boldsymbol{OK_{ex}}$	6.34E+04	2.95E+03	6.74E+04	3.34E+03	1.02E+05	1.25E+04	1.24E+04	1.20E+03
$\boldsymbol{OK_{Ga}}$	1.24E+06	1.73E+05	7.52E+04	3.84E+03	3.09E+09	1.16E+05	2.79E+08	1.05E+04
$\boldsymbol{OK_{St}}$	9.13E+04	3.80E+03	7.02E+04	3.46E+03	1.22E+05	1.48E+04	1.40E+04	1.40E+03

Table B.32 Cumulative production PLED 20 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	7.96E+04	4.23E+03	1.22E+05	5.39E+03	1.26E+05	1.82E+04	1.89E+04	1.77E+03
$\boldsymbol{OK_{Sp}}$	8.15E+04	4.36E+03	9.05E+04	5.04E+03	1.16E+05	1.73E+04	1.53E+04	1.68E+03
$\boldsymbol{OK_{ex}}$	7.93E+04	4.30E+03	8.61E+04	4.89E+03	1.19E+05	1.71E+04	1.51E+04	1.65E+03
$\boldsymbol{OK_{Ga}}$	5.37E+05	1.67E+05	1.03E+05	5.76E+03	3.95E+09	1.67E+05	3.56E+08	1.51E+04
$\boldsymbol{OK_{St}}$	1.02E+05	5.46E+03	9.00E+04	5.01E+03	1.45E+05	2.07E+04	1.71E+04	1.96E+03

Table B.33 Cumulative production PLED 25 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	9.46E+04	5.60E+03	1.46E+05	7.33E+03	1.41E+05	2.28E+04	2.20E+04	2.24E+03
$\boldsymbol{OK_{Sp}}$	9.72E+04	5.83E+03	1.10E+05	6.61E+03	1.30E+05	2.16E+04	1.79E+04	2.11E+03
$\boldsymbol{OK_{ex}}$	9.42E+04	5.73E+03	1.04E+05	6.55E+03	1.33E+05	2.13E+04	1.75E+04	2.08E+03
$\boldsymbol{OK_{Ga}}$	2.45E+05	1.46E+05	1.31E+05	7.93E+03	4.18E+09	2.46E+05	3.76E+08	2.22E+04
$\boldsymbol{OK_{St}}$	1.03E+05	7.06E+03	1.09E+05	6.70E+03	1.61E+05	2.61E+04	1.97E+04	2.49E+03

Table B.34 Cumulative production PLED 30 years

Scenario	All Coordinates		Cluster1		Cluster2		Sum of Clusters	
Method	μ	σ	μ	σ	μ	σ	μ	σ
$\boldsymbol{IDW}$	1.08E+05	6.94E+03	1.68E+05	9.26E+03	1.52E+05	2.69E+04	2.48E+04	2.67E+03
$\boldsymbol{OK_{Sp}}$	1.10E+05	7.34E+03	1.25E+05	8.28E+03	1.41E+05	2.52E+04	2.00E+04	2.49E+03
$\boldsymbol{OK_{ex}}$	1.08E+05	7.15E+03	1.21E+05	8.21E+03	1.43E+05	2.50E+04	1.97E+04	2.47E+03
$\boldsymbol{OK_{Ga}}$	2.30E+05	1.34E+05	1.58E+05	1.02E+04	4.26E+09	2.63E+05	3.84E+08	2.37E+04
$\boldsymbol{OK_{St}}$	1.10E+05	8.50E+03	1.27E+05	8.39E+03	1.72E+05	3.09E+04	2.20E+04	2.97E+03