

# Development of proxy models for petroleum reservoir simulation: a systematic literature review and state-of-the-art

Luciana Maria Da Silva<sup>1</sup>, Guilherme Daniel Avansi<sup>2</sup>, Denis José Schiozer<sup>3</sup>

<sup>1</sup>Department of Energy of the Faculty of Mechanical Engineering, University of Campinas, Brazil and Department of Mathematical Sciences, Durham University, UK

<sup>2,3</sup>Department of Energy of the Faculty of Mechanical Engineering, University of Campinas, Brazil

**Abstract**— Proxy models are derived mathematical functions developed as substitutes for reservoir flow simulators. Several types of proxy models are reported in the literature, for instance, response surface models, surrogate models, or metamodels. These models are fast methods, recommended for their efficient response time to approximate model responses and, therefore, useful in the decision-making process related to reservoir management. These studies focus on modelling a limited set of factors, applications, and case studies of any technique. A systematic literature review (SLR) is performed to gather the aspects prompting the modelling of proxy models in the literature and state-of-the-art. For this, a set of search keywords with appropriate string were utilised to extract the most important studies that satisfied all the criteria defined and classified under journal and conference paper categories. The papers were condensed after removing redundancy, repetition and similarity through a sequential and iterative process. From the analysis carried out, several gaps were identified, especially during the proxy model construction. Proxy models have already been discussed in petroleum engineering as a representation of the real system of reservoir flow simulator software. However, the proxy model response is faster but has yet to establish the issues of uncertainty in the outputs. There is a need for the integration of fast methods and reservoir simulators which can improve and accelerate results within acceptance criteria and accuracy in decision-making processes related to reservoir management.

**Keywords**— Petroleum Engineering, Proxy Model, Reservoir Simulator, State-of-the-art, Systematic Literature Review.

## I. INTRODUCTION

The decision analysis applied to the development and management of petroleum fields involves risk due to several uncertainties, mainly in the reservoir and fluid parameters, economic model, operational availability, and high computational cost. A new methodology based on 12 steps for integrated decision analysis considering reservoir simulation, risk analysis, history matching (HM), uncertainty reduction techniques, representative models, and selection of production strategy under uncertainty, which is necessary for the decision-making process was developed by [1]. The authors used a low-fidelity reservoir simulation model directly to predict field performance and quantify risk.

High (HFM), Medium (MFM), and Low (LFM) Fidelity Models assume reservoir conditions and

characteristics and physical laws (flows in porous media), while proxy models do not. HFM are models whose degrees of representativeness of geological, geophysical, fluid information, and recovery process are notable with high accuracy and precision. MFM are models whose geological, geophysical, fluid information and recovery processes have already undergone simplifications to reduce the degree of accuracy and computational time. These are used in production forecasting processes (mainly probabilistic) or those that demand hundreds and even thousands of simulations. LFM are models whose geological, geophysical, fluid information, and recovery processes have already undergone significant simplifications and their precision, accuracy and computational time are low. More details in [2].

A proxy model also called surrogate model, metamodel or response surface is a representation of a real system or

its simulations [3]. It becomes advantageous, especially when the direct evaluation of the system is either impossible or involves a high computational cost to simulate [4]. Therefore, a proxy model is considered to be an efficient substitute for the simulation tool at higher levels of reservoir study including uncertainty analysis, risk analysis and production optimisation [5], and also to elaborate the risk curves [6], especially time-consuming simulators [3]. In other words, in cases where proxy models can effectively represent important output parameters, they can be used as an adequate substitution for full reservoir simulators [7].

Proxy model constructions are held as mathematical derived functions, which imitate the output of a simulation model to selected input parameters [7]. According to the authors [6] and [8], if reservoir simulation studies were conducted with mathematical and statistical techniques, proxy models could estimate how the variation of input factors affects reservoir behaviour with a relatively small number of reservoir simulation models.

The purpose of the proxy models is to reduce the number of simulated models to evaluate a determining search space. It may lose a certain degree of accuracy due to the process of proxy modelling [9], but there is a reduction in computational time. Due to these reasons, obtaining an accurate proxy model is usually critical, and the model discrepancy has to be taken into account [10]. In petroleum exploration and production, the decision-making process, history matching, production strategy optimisation and economic evaluation of oil field must consider the risk involved through quantifying the impact of uncertainties on the performance of the petroleum field [6].

Numerous practical applications in uncertainty quantification, history matching, optimisation, and forecasting are increasingly involved in proxy modelling. The number and diversity of the proxy models development have widely increased as substitutes for reservoir flow simulators. On the other hand, a lack of better choice of the objective function and the methods able to correlate input and output are identified as the typical characteristics, which cause quality issues that might adversely influence the proxy models development.

Development of proxy models requires considering various factors, such as the selection of statistical and mathematical models, computational time, uncertainty quantification so forth. The initial knowledge on the effects of these factors on the development is fundamental to obtain an accurate model. Hence, a wide variety of proxy model application can be found in petroleum

engineering to investigate the effect of these factors on proxy modelling. However, each study investigates a limited set of particular input and, as a result, an extensive summary of existing literature on petroleum engineering is a valuable source for researchers in proxy model development.

This study aims to present the aspects identified in the studies analysed and thus present the current state of the research. A systematic literature review (SLR) is performed to gather the elements prompting the modelling of proxy models in the literature and state-of-the-art in petroleum reservoir engineering. For this, a set of search keywords with the appropriate string were utilised to extract most important studies that satisfied all the criteria defined in the relation between proxy model developments and classified under journal and conference paper categories. The information obtained in SLR and state-of-the-art is useful for industry experts and researchers.

This paper is structured as follows: Section II presents the background studies of the proxy model; Section III provides an overview of research methodology; Section IV summarises the results, which were essential to answer our research questions; Section V highlights the discussion showing the gaps we identified for future research and the state-of-the-art; Section VI presents the conclusion of the paper.

## II. BACKGROUND STUDIES OF PROXY MODEL

There were no systematic reviews that originated under the modelling of proxy models or aspects in petroleum engineering. From the literature gathered, the authors searched and examined the studies performed between the years 2007 and 2017 in digital libraries to develop the SLR. Still, we do not limit to this years to the state-of-the-art aspects showing aspects until 2020.

Development of proxy models has been performed on various models for reservoir flow simulation, which can be used for forecasting, optimisation of production, history matching, characterisation of reservoir properties, uncertainty and risk analysis, and production strategy selection. These proxy models can be polynomial regression models, ordinary kriging models, artificial neural networks (ANNs), and radial basis functions (RBFs), response surface methodology (RSM), design of experiment (DE), and other.

We can find in the literature a wide range of proxy model development for application in petroleum engineering, for example, a new approach to improve Bayesian HM [11, 12]. The authors [13] integrated a framework for field-scale modelling, HM, and robust

optimisation of field scale low salinity waterflooding (LSW). An approach using the SRM for optimisation [14-19]. The authors [20] addressed the decision-making process over the determination of oil & gas production strategies.

Some papers applied ensemble Kalman Filter (EnKF) with an objective, for example, the authors [21-24] for the analysis of uncertainty quantification and optimisation method, [25, 26] to automatise HM, [27] for estimation of channel permeability in a bimodal distribution, and [28] for the integration of well-test data into heterogeneous reservoir models. The authors [29] combined EnKF with Markov Chain Monte Carlo (MCMC) to obtain a more accurate characterisation of uncertainty; [30] combined EnKF with genetic algorithm.

The authors [31] made the comparison of SRM with least square support vector machine. Use of experimental design to develop response surface [32-41], integrated with Monte Carlo simulations to characterise the response surface and to estimate the uncertainty [42, 43]. Application of Bayesian multi-stage MCMC approach, based on an approximation with a linear expansion to reduce high computational costs [44], more accurately obtained model uncertainty and also assists in production-forecast business decisions [45], with Bayesian workflow based on two-step MCMC inversion [46].

In [47] was presented a method to select a subset of reservoir model computing the statistics (P10, P50, P90) of the response of interest; use of the genetic algorithm to improve the process of optimisation [48]. Application of an approach with fuzzy analytical hierarchy process for compositional simulation studies of the CO<sub>2</sub> injection

process [49]. The authors [50] developed a semi-analytical fast model for optimal field development strategy. The authors [51] used principal component analysis (PCA) and elastic gridding. Application of a robust reservoir simulator with the application of kriging models [10; 52, 53]; in a closed-loop [54, 55]. Combination of Karhunen-Loeve (KL) expansion and probabilistic collocation method for uncertainty analysis [56]. Development of an emulator utilised Bayes Linear [23, 57]; development of a proxy model to predict cumulative oil production and steam injection profiles [58].

The authors [7, 59] proposed the application of polynomial chaos proxy efficiently sample with MCMC and ANNs, respectively. Application of ANNs in the form of gene expression programming is applied through an extensive statistical manner [60] in HM [61-63]. In recent years, ANN training has been accomplished to identify the non-linear relationships between various input and output variables [3, 5; 64-69] used ANNs integrated to polynomial regression for risk analysis and forecasting.

### III. PROCEDURE FOR SYSTEMATIC LITERATURE REVIEW

SLR is the best method available to generate scientific evidence based on the summary of the significant publications concerning a specific topic or research question [70]. Due to this, the methodology was undertaken based on [71] to survey the existing knowledge about the development of proxy models for petroleum reservoir simulation. The SLR process applied can be seen in Fig. 1.

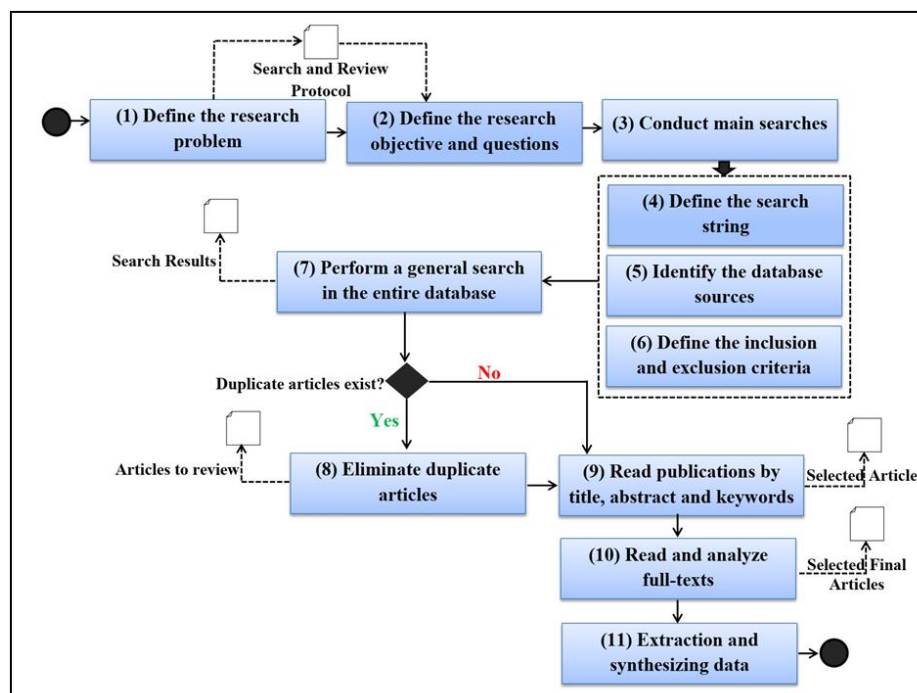


Fig. 1 Workflow with each step achieved in SLR

The authors performed the planning of the review, from which the research problem, objective and questions were defined (steps 1 and 2). Therefore, we obtained the search and review protocols. Afterwards, we performed the definition of the primary searches (step 3) based on search string (step 4), database sources (step 5), inclusion and exclusion criteria (step 6), resulting in the general search in the entire database (step 7). From the results of the examination, the duplicate articles were eliminated (step 8), obtaining a list of selected papers which were read by title, abstracts, and keywords (step 9). After the partial reading, we got a list with the selected final articles which were thoroughly read and analysed (step 10).

We specified the details of the SLR methodology in the following subsections: research questions, search and review protocol, define the search string, identify the database sources, and define the inclusion and exclusion criteria. The extraction and synthesising concerning the general search in the entire database, numbers of eliminated duplicate articles, numbers and criteria of reading publications by title, abstract and keywords, and numbers of reading and analysed full-texts are in Section 4 (step 11).

### 3.1 Research problem and questions

The identification of the aspects of proxy model development requires a clear and explicit analysis of the research problem and theoretical concept (step 1). From

this, we formulated research questions for this SLR (step 2):

RQ1: How many proxy model studies have been performed from 2007 to 2017?

RQ2: What were the research topics addressed to the publication?

RQ3: What were the problems investigated and presented in the literature to the development of the proxy model?

RQ4: Why use the proxy model?

About RQ1, we identified that the term “systematic literature review” was not in common usage in the petroleum area. In contrast, in Information and Software Technology, Chemistry, Business Administration and Medicine, it is diffused. The authors [71] highlight that there are rigorous example literature reviews before 2004 in the software engineering area. Therefore, based on RQ1, we identified the number of articles published per year, the journals, conferences, and database which published about the development of proxy models. Concerning RQ2 the aspects of the petroleum engineering topic area and the model-based decision were considered (closed-loop reservoir development and management – CLRDM) developed by [1]. For RQ3 the problems in the decision-making process for petroleum reservoir simulation related to the CLRDM model were considered, such as, overcome computational costs, computational time demand and performance of a reservoir simulator, reduced human resources and fidelity model. In RQ4, we considered the

proxy models and emulators identified during the reading of articles.

### 3.2 Search and Review Protocol

A search and review Protocol is essential in all SLR to guarantee the efficiency of the selected studies. For this, it is necessary to define the research problem in parallel with the research objectives and questions, as shown in Fig. 1 (steps 1 and 2).

The protocol for this review depended on step 3 being developed in three stages (from step 4 to step 6): P1: Define the search string, P2: Select the literature database, and P3: Define the inclusion and exclusion criteria, this defines the protocol that was used to perform the search in the sources defined, which will be explained in the subsections: define the search string, identify the database sources, and define the inclusion and exclusion criteria.

### 3.3 Define the search string

SLR is a known technique for reviewing the literature with vast search information of the subject in the discussion from all relevant sources. Due to this, a systemic method to formulate search keywords was defined, considering the following issues:

- Setting of significant terms based on the research question;
- Setting of similar words for significant terms;
- Setting of relevant keywords in any applicable studies;
- Using Boolean operators “OR” and “AND” as an alternative to linking terms.

We defined the search string with focus on related studies of petroleum simulator and proxy model, i.e., an exact string “(“oil” OR “petroleum”) AND “uncertainty” AND “simulator””. The first part of the string was the focus area of the research. We included the words “*uncertainty*” and “*simulator*” to disqualify studies which are related to fields different from petroleum engineering.

The authors opted not to utilise the words proxy model as the exact phrase since, in most of the search queries, there are numerous studies in which proxy models are related as surrogate, metamodel or response surface. If “proxy model” had been utilised alone, the search would lose significant results that use the terms: surrogate, metamodel or response surface.

### 3.4 Identify the database sources

To perform the SLR and to find the relevant studies, we searched the following seven major electronic libraries, six general and one specific to the area of petroleum engineering.

- ACM Digital Library (<http://dl.acm.org>)
- IEEE Xplore (<http://ieeexplore.ieee.org>)
- ScienceDirect (<http://www.sciencedirect.com>)
- Scopus (<http://www.scopus.com>)
- SpringerLink (<http://link.springer.com>)
- Web of Science (<http://apps.webofknowledge.com>)
- OnePetro (<https://www.onepetro.org>)

In this research, we did not select the papers manually, and for this selection, we used an automatic selection criteria (scripts in Python language) developed by [72].

### 3.5 Define the inclusion and exclusion criteria

The definition of the inclusion and exclusion criteria was based on the determination of an objective and question research. We applied the inclusion and exclusion criteria in the resulting publications, after eliminating the duplicated articles and identifying which would be relevant to this SLR. Table 1 shows the inclusion and exclusion criteria considered in the database source.

We initially applied the inclusion and exclusion criteria in the entire database (step 7). The first criterion considered were articles in the English language, published from 2007 to 2017, peer-reviewed publications and whether their abstract contained any word of the string. After the search finished generating the list of articles, we used the string to analyse the full papers. If at least one term of the string had an association with the title, keywords and abstract, we included the article in the significant study list. For duplicated articles in multiple databases, we removed them and used one copy in the analysis (step 8). After, in step 9, in the inclusion and exclusion criteria process, we read the title, abstract and keywords to applicate the five assessments (Table 2). We generated these assessments to analyse the applicability and development of articles as exclusion criteria.

Table 1: Inclusion and exclusion criteria for the analysis of articles selected in the database.

Considered Criteria	
Inclusion	Exclusion
Period of publication from 1 January 2007 to 31 December 2017	Duplicated publications of the same study in more than one database



Publications published in the English language	Non-English Language publication
Publications that were peer-reviewed	Publications without bibliographic information
Publications which address proxy model and reservoir flow simulators software	Publications which do not address proxy model or only include reservoir flow simulator software
Publications that focus on the development of the proxy model	Publications that only identify the technological aspects of the tools used
Publications that presented the keywords which belong to the string determined in this SLR	Publications that do not present the keywords which belong to the string determined in this SLR
Journal with Scimago (SJR) $\geq 0.2$ or JCR $\geq 0.5$ and Conference (peer-reviewed)	Other knowledge of the area

Table 2: Five assessments utilized for partial analysis of the articles.

Assessment	Description
1	The articles address reservoir characterization and/or uncertainty and/or optimization and/or risk and/or history matching and/or forecasting, it works with reservoir simulator software, but it did not develop a proxy model or apply
2	The articles were applied in another area of knowledge, or they only mentioned reservoir simulator software
3	Revision article: present difficulties to be reproduced, being applied to specific parameters without a new technique development
4	Description of the combination of techniques in oil reservoir with reservoir simulator software
5	Identify the technological aspects of tools used

As an initial step, a general search was made, which was inside the inclusion criteria but was outside the scope of five assessments. It is essential to highlight that; this application is to analyse the significant researches which will be adequate to answer all RQs. Subsequently, we excluded various papers. And we selected 117 articles to read them thoroughly.

In step 10, the full reading of the selected articles, we generated nine assessment questions for data extraction, from QE1 to QE8. An assessment question “Yes(Y)” = 1, “Partly(P)” = 0.5, “No(N)” = 0 or “Unidentified (U)” was also included to evaluate the contribution of each article during the proxy definition and construction. Besides, some articles may have a more straightforward proxy model development, focusing on application without many details and, because of this, various papers were considered unrelated to the development of proxy models, after reading the full article.

QE1: What was the method used for data sampling?

QE2: What was the type of proxy model performed?

QE3: What was the objective function used?

QE4: Was there any performance addressed to computational time?

QE5: What were the aspects additionally addressed in the article?

QE6: What were the problems presented in the article?

QE7: What was the focus of the article analyzed?

QE8: Was there any article relevant to the development or application of proxy models?

Concerning QE1, when the method used for data sampling is explicitly defined (Y), it is implicit (P), or it is not defined or cannot be readily explicit (N). For QE2, when the proxy models performed are explicitly (Y), they are implicit (P), or they are not defined or cannot be readily explicit (N). About QE3, if the objective function is explicitly defined (Y); it is implicit (P); it is not defined or cannot be expressly identified (N). For QE4, if the performance addressed was defined for proxy model development or applied the modelling proposed (Y), it was defined for reservoir numerical simulator (P), or it was not implemented (N). Concerning QE5, the additional aspects are explicitly described (Y); they are implicit (P), or they cannot be expressly identified (N). For QE6, the problems presented are explicitly defined (Y); they are implicit (P), or they are not or cannot be expressly specified (N). For QE7, article approached modelling or experiment of the proxy model (Y); it was an application, literature review or technique (P); the paper analysed cannot be explicitly

identified (N). For QE8, the article approached obtained a score of >4.0 (Y); it got a score of ≤ 4 (N). For all questions, we considered (U) in case the information not specified. Table 3 presents the keywords considered in the article as an answer to all questions.

#### IV. RESULTS

This section presents the results (step 11), which we divided into three parts: perform a general search in the entire database (step 7) and the article selection process (steps 8 and 9); results from article reading and classification (step 10); quality factors.

4.1 Perform a general search in the entire database and the article selection process

We developed an SLR to gather the aspects prompting proxy model development in the literature. For this, we utilised a set of search keywords with appropriate string to extract the essential researches that satisfied all the criteria defined and classified under journal and conference paper categories, in seven scientific electronic library databases, resulting in 4,687 publications from January 2007 to December 2017. We showed the distribution of articles and the types (Journal and Conference) in each database in Fig. 2.

Table 3: Defined answer for application in reading the final articles

QE1	QE2	QE3	QE4	QE5	QE6	QE7	QE8
Random	Multivariate Kriging	Np	Applied in metamodel developed	Uncertainty analysis	Computational Time	Literature Review	Yes
Stratified	Artificial Neural Network	Wp	Applied in a simulator used	History Matching	Computational resource	Application	No
Systematic	Response Surface/ Surrogate	NPV	Applied the modelling proposed	Reservoir Characterization	Type of data	Technique	-
Cluster	Fuzzy Logic	ROI	No measurement implemented	Optimization	Unidentified	Modelling	-
Rank	Bayesian	Capillary pressure	Unidentified	Production Strategy Selection	-	Experimental	-
Unidentified	Kalman Filter	Others	-	Risk Analysis	-	-	-
-	Experimental Design	Unidentified	-	Unidentified	-	-	-
-	Other metamodel	-	-	-	-	-	-
-	Unidentified	-	-	-	-	-	-

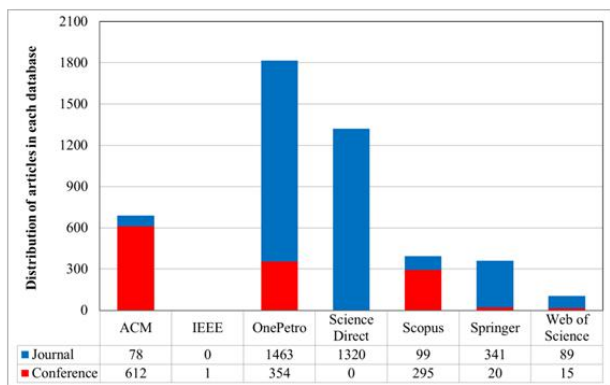


Fig. 2 Results obtained from string application in the databases; in red - the conference numbers and in blue - the journal numbers

It is possible to see the results of the selected articles in Fig. 2, a total of 4,687 papers, where 3,390 were published in journals and 1,297 in conferences. Fig. 3 shows the results of each step of article selection and the percentages of each publication. From the 4,687 publications obtained, we applied the exclusion and inclusion criteria process and resulted in 317 usability publications (in blue), which represents 6.76% of the selected publications per database. We applied a sequential and iterative approach (python script), and we condensed the publication removing redundancy, repetition and similarity (in red), which represented 31.55%. The publications excluded they were in multiple databases. We reduced the publications for the reading of title, abstract and keywords, and after that, we removed 100 publications (46.08%) based on the five assessments shown in Table 2. Finally, we obtained 117 papers to read them thoroughly, representing 53.92% (in green).

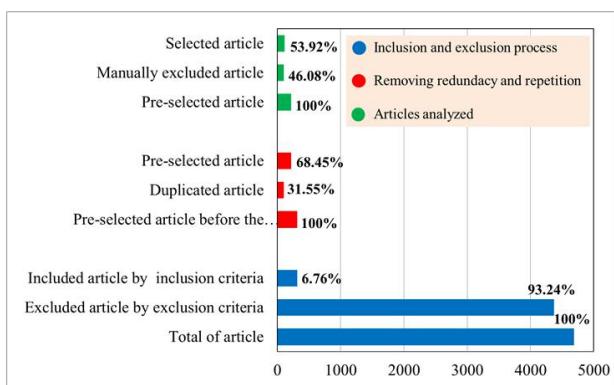


Fig. 3 The number of searches in the database in each step. In blue- the inclusion and exclusion criteria process; in red- the phase after removing redundancy and repetition; in green- the articles analysed.

Concerning the final process of selecting the publications, we initially worked with seven databases. In the chosen article process, only four databases returned publications. OnePetro electronic library produced the highest number of publications (full reading). Fig. 4 shows the distribution of pre-selected and selected publications over the years, which returned the string application in this SLR.

Fig. 4 shows the distribution of publications per year; the blue axis shows the quantity of pre-selected publications, and the red axis indicates the number of selected publications. About the selected publications, it is possible to observe that the years 2008 and 2014 presented the highest number of publications. In analyzing the numbers obtained in 2008, 10 publications were in conferences while seven publications were in journals. In contrast, 12 and 7 publications were published in conferences and journals in 2014, respectively. We noticed that in 2017, one paper was obtained from the conference while eight publications in journals. Other reasons for the changes over the years, we considered only peer-reviewed publications, and the journal must have  $SJR \geq 0.2$  or  $JCR \geq 0.5$  and focus on the development of proxy models in the petroleum engineering area. We analysed the barrel price of crude oil (Brent) in dollars [73] and observed that when the publication numbers increased, the price per barrel reduced. We performed the Pearson correlation with a 5% significance level. The Pearson correlation between the cost of each barrel and the selected article numbers was  $-0.65$  ( $p < 0.031$ ).

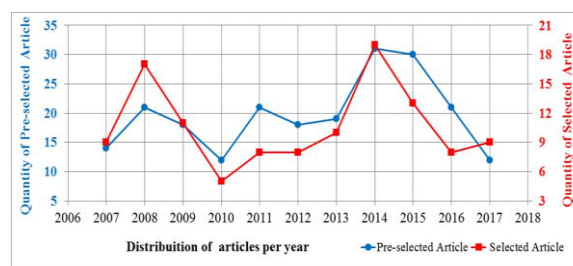


Fig. 4 Distribution of articles per year in the process of selection. In the blue - pre-selected article process and, in the red - selected article process.

#### 4.2 Results of reading and classification of articles

The authors developed nine assessment questions and defined answer for application in reading the final articles presented in Section 3.5. Table 4 shows the classification of the 117 publications selected with the percentage-based in each assessment question (Section 3.5).

Table 4 presents the classification of the 117 publications selected about each assessment question



(Table 3). Concerning QE1, the method most used for data sampling is a random sample, at 63.25%. To QE2, we observed that each ANN and RSM presented 11.96% in proxy model development. QE3 showed that the most utilized objective function was the NPV, at 23.08%. Concerning QE4, the authors used more performance on the modelling proposed. About QE5, which refers to the additional aspects, the one most used was “optimization” and “history matching,” which are essential parts of a reservoir process that highly need proxy models. QE6, the most detected problem is of computational time, while

QE7 shows that the focus of the article is mostly on “application”, at 47.01% of publications.

Table 5 shows the results obtained from 40 publications selected for full reading by score obtained, and the papers presented only study application or technique application. In some cases, it was not possible to identify the procedure used to model the proxy, totalising 34.19% of 117 publications selected. In a total of 32 publications, it was not possible to identify the modelling on the proxy model.

Table 4: Classification of the 117 publications selected concerning each assessment question

QE	Assessment question for data extracted	Answer	Quantity	(%)
1	What was the method used for data sampling?	Random	74	63.25
		Stratified	5	4.27
		Systematic	2	1.71
		Cluster	3	2.56
		Rank	9	7.69
		Unidentified	24	20.52
2	What was the type of proxy model performed?	Multivariate Kriging	6	5.13
		Artificial Neural Network	14	11.96
		Response Surface/Surrogate	14	11.96
		Fuzzy Logic	2	1.71
		Bayesian	8	6.84
		Kalman Filter	10	8.55
		Experimental Design	6	5.13
		Other metamodels	31	26.50
Unidentified	26	22.22		
3	What was the objective function used?	Np	6	5.13
		Wp	2	1.71
		NPV	27	23.08
		ROI	1	0.86
		Capillary pressure	3	2.56
		Others	41	35.04
		Unidentified	37	31.62
4	Was there any performance addressed to computational time?	Applied in metamodel developed	18	15.38
		Applied in a simulator used	6	5.13
		Applied the modelling proposed	56	47.86
		No measurement	8	6.84
		Unidentified	29	24.79

5	What were the aspects additionally addressed in the article?	Uncertainty analysis	14	11.97
		History Matching	29	24.78
		Reservoir Characterization	10	8.55
		Optimization	31	26.49
		Production Strategy Selection	5	4.27
		Risk Analysis	14	11.97
		Unidentified	14	11.97
6	What were the problems presented in the article?	Computational Time	76	64.96
		Computational resource	9	7.70
		Type of data	16	13.67
		Unidentified	16	13.67
7	What was the focus of the article analyzed?	Literature Review	1	0.86
		Application	68	58.12
		Technique	3	2.56
		Modelling	22	18.80
		Experimental	23	19.66

Table 5: Result quality scores of selected publications with a score of  $\leq 4.0$

Number	Publication	QE1	QE2	QE3	QE4	QE5	QE6	QE7	QE8(Score)
1	[74]	0.0	0.0	0.0	0.0	0.0	1.0	0.5	1.5
2	[75]	0.0	0.0	1.0	0.0	0.0	0.0	0.5	1.5
3	[76]	0.0	1.0	0.0	0.0	0.0	0.0	0.5	1.5
4	[77]	0.0	0.0	0.0	0.0	1.0	0.0	0.5	1.5
5	[78]	0.0	0.0	1.0	0.0	0.0	0.0	1.0	2.0
6	[79]	0.5	0.0	0.0	1.0	0.0	0.0	1.0	2.5
7	[80]	0.0	0.0	0.0	0.0	1.0	1.0	0.5	2.5
8	[81]	0.0	1.0	0.0	0.0	1.0	0.0	0.5	2.5
9	[82]	0.0	0.0	1.0	0.0	0.0	1.0	0.5	2.5
10	[83]	0.0	1.0	0.0	0.0	0.0	1.0	0.5	2.5
11	[84]	0.5	0.0	0.0	0.5	0.0	1.0	0.5	2.5
12	[85]	0.0	0.0	0.0	1.0	0.0	1.0	0.5	2.5
13	[86]	0.0	0.0	0.0	0.0	1.0	1.0	0.5	2.5
14	[87]	1.0	0.0	0.5	0.0	1.0	0.0	0.5	3.0
15	[88]	0.5	0.0	0.0	1.0	1.0	0.0	0.5	3.0
16	[89]	0.5	0.0	1.0	0.0	0.0	1.0	0.5	3.0
17	[90]	0.0	0.0	0.5	0.0	1.0	1.0	0.5	3.0
18	[91]	0.5	0.0	0.0	0.0	1.0	1.0	0.5	3.0
19	[92]	0.0	1.0	0.0	0.0	1.0	0.0	1.0	3.0
20	[93]	0.0	0.0	0.5	0.0	1.0	1.0	0.5	3.0

21	[94]	0.5	0.0	0.0	1.0	1.0	0.0	0.5	3.0
22	[95]	0.5	1.0	0.0	0.0	1.0	0.0	0.5	3.0
23	[96]	0.5	0.0	0.5	0.0	1.0	1.0	0.5	3.5
24	[97]	1.0	1.0	0.0	0.0	0.0	1.0	0.5	3.5
25	[98]	0.5	0.0	0.5	0.0	1.0	1.0	0.5	3.5
26	[99]	0.0	0.0	1.0	0.0	1.0	1.0	0.5	3.5
27	[100]	0.0	1.0	0.0	0.0	1.0	1.0	0.5	3.5
28	[101]	0.0	0.0	1.0	0.0	1.0	1.0	0.5	3.5
29	[102]	0.0	0.0	1.0	1.0	1.0	0.0	0.5	3.5
30	[103]	0.0	0.0	0.0	1.0	1.0	1.0	0.5	3.5
31	[104]	1.0	0.0	0.0	0.0	1.0	1.0	0.5	3.5
32	[105]	1.0	0.0	0.0	1.0	1.0	0.0	0.5	3.5
33	[106]	0.5	0.0	0.0	1.0	1.0	1.0	0.5	4.0
34	[107]	0.5	0.0	1.0	0.0	1.0	1.0	0.5	4.0
35	[108]	0.5	0.0	1.0	0.0	1.0	1.0	0.5	4.0
36	[109]	0.5	0.0	1.0	0.0	1.0	1.0	0.5	4.0
37	[110]	0.0	1.0	0.5	1.0	0.0	1.0	0.5	4.0
38	[111]	0.5	0.0	1.0	0.0	1.0	1.0	0.5	4.0
39	[112]	0.0	0.0	0.0	1.0	1.0	1.0	1.0	4.0
40	[113]	0.0	0.0	1.0	0.5	1.0	1.0	0.5	4.0

Table 6 shows the result quality scores of selected publications with a score of > 4.0. We identified a total of 78 publications as having a real contribution to the definition of a proxy, and the construction method of the

proxy used, totalising 65.81% of the 117 publications selected based on our criteria.

Table 6: Result quality scores of selected publications with a score of > 4.0

Number	Publication	QE1	QE2	QE3	QE4	QE5	QE6	QE7	QE8(Score)
1	[114]	1.0	0.0	1.0	0.0	1.0	1.0	0.5	4.5
2	[115]	1.0	0.0	0.0	1.0	1.0	1.0	0.5	4.5
3	[116]	1.0	1.0	0.0	1.0	0.0	1.0	0.5	4.5
4	[117]	0.5	0.0	1.0	1.0	1.0	0.0	1.0	4.5
5	[118]	0.5	0.0	0.5	1.0	1.0	1.0	0.5	4.5
6	[58]	0.5	1.0	0.5	1.0	1.0	1.0	0.5	4.5
7	[119]	0.5	1.0	0.5	1.0	1.0	0.0	0.5	4.5
8	[120]	0.0	1.0	1.0	0.0	1.0	1.0	0.5	4.5
9	[25]	0.5	1.0	0.0	1.0	1.0	1.0	0.5	5.0
10	[121]	1.0	0.0	0.5	1.0	1.0	1.0	0.5	5.0
11	[122]	1.0	0.0	0.0	1.0	1.0	1.0	1.0	5.0
12	[123]	0.0	1.0	0.0	1.0	1.0	1.0	1.0	5.0
13	[38]	0.5	1.0	0.0	1.0	1.0	1.0	0.5	5.0

14	[50]	0.5	1.0	1.0	1.0	1.0	0.0	0.5	5.0
15	[69]	0.5	1.0	0.5	1.0	0.0	1.0	1.0	5.0
16	[15]	1.0	1.0	1.0	1.0	0.0	1.0	0.5	5.5
17	[22]	1.0	1.0	0.0	1.0	1.0	1.0	0.5	5.5
18	[23]	1.0	1.0	1.0	0.0	1.0	1.0	0.5	5.5
19	[26]	1.0	1.0	0.0	1.0	1.0	1.0	0.5	5.5
20	[34]	1.0	1.0	0.0	1.0	1.0	1.0	0.5	5.5
21	[40]	1.0	1.0	1.0	1.0	1.0	0.0	0.5	5.5
22	[59]	1.0	1.0	0.0	1.0	1.0	1.0	0.5	5.5
23	[124]	1.0	1.0	0.5	1.0	0.5	1.0	0.5	5.5
24	[17]	0.5	1.0	0.5	1.0	1.0	1.0	0.5	5.5
25	[43]	0.5	1.0	0.0	1.0	1.0	1.0	1.0	5.5
26	[44]	0.5	1.0	0.5	1.0	1.0	1.0	0.5	5.5
27	[45]	0.5	1.0	0.5	1.0	1.0	1.0	0.5	5.5
28	[51]	0.5	1.0	0.0	1.0	1.0	1.0	1.0	5.5
29	[21]	1.0	1.0	0.5	0.0	1.0	1.0	1.0	5.5
30	[28]	1.0	1.0	1.0	0.0	1.0	1.0	0.5	5.5
31	[68]	1.0	1.0	0.0	1.0	1.0	1.0	0.5	5.5
32	[5]	1.0	1.0	0.0	1.0	1.0	1.0	1.0	6.0
33	[62]	1.0	1.0	0.5	1.0	1.0	1.0	0.5	6.0
34	[64]	1.0	1.0	0.5	1.0	1.0	1.0	0.5	6.0
35	[11]	0.5	1.0	0.5	1.0	1.0	1.0	1.0	6.0
36	[16]	0.5	1.0	1.0	1.0	1.0	1.0	0.5	6.0
37	[65]	0.5	1.0	0.5	1.0	1.0	1.0	1.0	6.0
38	[125]	0.5	1.0	1.0	1.0	1.0	1.0	0.5	6.0
39	[20]	1.0	1.0	1.0	0.0	1.0	1.0	1.0	6.0
40	[47]	1.0	1.0	0.5	1.0	1.0	1.0	0.5	6.0
41	[57]	1.0	1.0	0.0	1.0	1.0	1.0	1.0	6.0
42	[61]	1.0	1.0	0.5	1.0	1.0	1.0	0.5	6.0
43	[66]	1.0	1.0	0.5	1.0	1.0	1.0	0.5	6.0
44	[7]	0.5	1.0	0.5	1.0	1.0	1.0	1.0	6.0
45	[35]	0.5	1.0	1.0	1.0	1.0	1.0	0.5	6.0
46	[36]	0.5	1.0	1.0	1.0	1.0	1.0	0.5	6.0
47	[37]	0.5	1.0	0.5	1.0	1.0	1.0	1.0	6.0
48	[3]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
49	[6]	1.0	1.0	1.0	1.0	1.0	1.0	0.5	6.5
50	[10]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
51	[12]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
52	[24]	1.0	1.0	1.0	1.0	1.0	1.0	0.5	6.5

53	[29]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
54	[30]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
55	[32]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
56	[46]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
57	[49]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
58	[42]	0.5	1.0	1.0	1.0	1.0	1.0	1.0	6.5
59	[55]	0.5	1.0	1.0	1.0	1.0	1.0	1.0	6.5
60	[9]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
61	[14]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
62	[31]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
63	[41]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
64	[52]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
65	[54]	1.0	1.0	1.0	1.0	1.0	1.0	0.5	6.5
66	[60]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
67	[67]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
68	[126]	1.0	1.0	0.5	1.0	1.0	1.0	1.0	6.5
69	[18]	0.5	1.0	1.0	1.0	1.0	1.0	1.0	6.5
70	[39]	0.5	1.0	1.0	1.0	1.0	1.0	1.0	6.5
71	[19]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0
72	[48]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0
73	[53]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0
74	[56]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0
75	[13]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0
76	[27]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0
77	[63]	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.0

We presented in Table 6 the results obtained from 77 publications selected for full reading, by score obtained, and the construction of the proxy model, where it is possible to identify the modelling or experiment developed.

#### 4.3 Quality factors

According to [71], SLRs are literature surveys with defined research questions, search process, data extraction and data presentation, whether the researchers referred to their study as a systematic literature review. Due to this,

we analysed the relationship between the score obtained with the QEs and the date of publication. In this analysis, we deemed the 77 relevant publications to the proxy model development. The average quality scores for publications considered as a contribution in the definition of a proxy model for each year is shown in Table 7.

Table 7: Analysis of quality scores for 77 publications considered relevant in proxy model development

	Years										
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Number of publications	7	12	9	2	4	5	6	13	7	5	7



Mean	5.79	5.88	5.94	6.25	6.13	5.40	5.67	5.85	5.86	6.10	6.21
Standard deviation	0.52	0.62	0.76	0.25	0.65	0.73	0.62	0.72	0.83	0.80	0.80

Table 7 indicates that for the years 2008 and 2014 have had relatively more publications based on our criteria.

## V. DISCUSSION

This section, we present the answers to our questions (Topic 3.1), which reported what has been investigated in the literature and considered in proxy model development.

5.1 How many proxy model studies were performed from 2007 to 2017?

Overall, we identified 117 publications. We extracted data and synthesised them to answer our research questions. We selected 77 publications which they were more relevant because the score obtained with the application of our research question was higher to 4.0. A total of 40 papers we considered less relevant because their application was simple, or it was not possible to identify the proxy model development or the modelling applied.

About analyse the proxy model performed in the literature, we identified six types of proxy models that are more utilised in the publications, others were also identified, and then an “other metamodel” class was created. This class represents 31% of the 117 publications which developed another type of metamodel that is different from the traditional one. It is possible to affirm from the literature that the proxy model is also identified as a surrogate, response surface methodology or metamodel, and emulator. Concerning the objective function used, we analyse 117 publications, and 35.04% used implicit objective functions while 31.62% did not define or it was not explicit. The greater focus of published articles was on “application”, some very detailed and some simple.

This SLR identified 52 articles published in journals, totalising 44.44% used to develop this research. Of these 52 publications published, 23.08% - SPE Journal; 21.15% - Journal of Petroleum Science and Engineering; 9.62% - Journal of Natural Gas Science and Engineering; 7.69% - Journal of Canadian Petroleum Technology; 5.77% - Petroleum Science and Technology; 3.85% - SPE Reservoir Evaluation & Engineering, and 28.85% are distributed in another 14 journals. About the conferences, 65 articles are published in 27 different Annals, totalising 56.56% of the 117 publications. Of these 65 publications, we noticed 18.46% in Proceedings - SPE Annual Technical Conference and Exhibition. Still, if we consider all the conferences organised by the Society of Petroleum

Engineering (SPE), they summarise to 61.54%. From the 27 Annals, we observed 17 organised by SPE. The conferences of ECMOR - European Conference on the Mathematics of Oil Recovery, IPTC - International Petroleum Technology Conference and SPE Canada Heavy Oil Technical Conference correspond to 7.69% each.

5.2 What were the research topics addressed to the publication?

Concerning the subject of the articles, six were related to research trends which belong the tree main petroleum areas: Past, Future (Decision-making) and Future (Reservoir Behavior, Production Forecast), as addressed in the model based on CLRDM by [1]. Fig. 5 illustrates the six topics (3 past factors; 2 future (decision-making); and one future (reservoir behaviour, production forecast)) approached in the publications, we identified them in different colours.

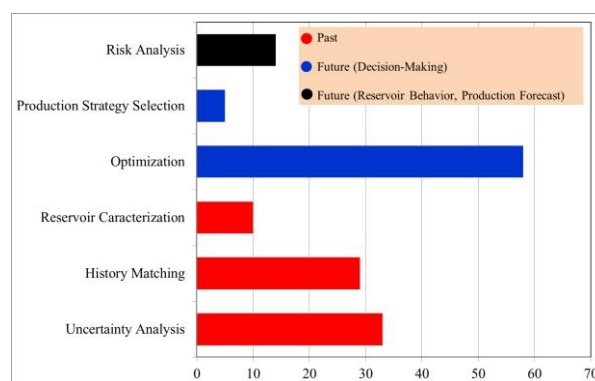


Fig. 5 The three main areas of petroleum reservoir studies related to the development and management of petroleum fields. In red, past; in blue, decision-making (future); in black, reservoir behaviour, production forecast (future).

For past (red) highlighting: uncertainty analysis, history matching, and reservoir characterization. In terms of future (blue) highlighting the aspects that addressed the decision-making process: optimization and production strategy selection. And for future (black) highlighting the elements that addressed reservoir behaviour and production forecast: risk analysis. In the 117 publications, only 14, or 11.97%, were not possible to identify the corresponding area.

It is essential to mention that there are several ways to classify uncertainties in reservoir simulation and

characterization. According to [1] we have: (1) geostatistical realizations of facies, porosity, NTG and permeability; (2) attributes: water relative permeability (krw), PVT, water-oil contact depth (WOC), rock compressibility (CPOR) and vertical permeability multiplier (kz); and (3) economic uncertainties (e.g. market values, taxes, costs and investments). Fig. 5 shows that the term “optimization” was present in most publications. From the data obtained from the production optimisation process, data are generated for running past information, and they are utilised in the future in the production strategy selection process.

There are publications which addressed the production data optimisation, optimisation integrated with uncertainty analysis; risk analysis; history matching; production strategy selection. In some publications the term “optimisation” can be combined to more than one word, for example: “optimisation”, “uncertainty analysis”, “history matching” and “reservoir characterisation”; or “optimisation”, “risk analysis” and “production strategy selection”; or “optimisation”, “uncertainty” and “risk analysis”. Another factor observed in Fig. 5 is the fact that only five publications of the 117 focused on the term “production strategy selection”. This term is essential in future decision-making processes because the development and management of petroleum fields involve risk due to several uncertainties. The authors [1] presented the integration of these six topics step by step with characterisation, long term production data, decision-making process, history matching, details, particularities and complexities.

### 5.3 What are the problems investigated and presented in the literature for the development of the proxy model?

Concerning the six topics shown in Fig. 5 obtained from SLRs are limited in decision-making, a large number of publications (76 papers) related to computational time as an essential factor in proxy model development. When the proposed proxy model dramatically reduces the computation time, it potentially carries out frequent execution of uncertainty quantification, history matching, risk analysis, and optimisation, resulting in efficient reservoir management and significant computational time reduction. For example: [7, 59] developed a proxy model using Polynomial Chaos Expansion to improve computational time when utilising numeric reservoir simulator. They obtained significant monetary benefits and computational time reduction.

Research on building a proxy model shows that there are critical problems with its development and accuracy. Among other issues, we identified the followings: high computational costs, computational times and performance

of reservoir simulator. Therefore, proxy models should consider development as an essential quality attribute to be achieved, because proxy models do not assume reservoir conditions and characteristics, and physical laws, enabling reduced computational time, reservoir simulator use and human resources.

The development of a proxy model requires considering various factors such as the size and complexity of the model. Knowledge of the effects of these factors in the six topics highlighted in Fig. 5 is essential both for research and practice. Hence, several publications have been performed to investigate the effect of these factors. In 76 publications (from a total of 117), the authors highlighted the importance of computational time reduction; 9 publications highlighted computational resource reduction; and 16 publications highlighted the type of data as an essential factor to be investigated in proxy model development. In another 16 publications, it was not possible to identify the problem present in proxy model development.

Each publication explores a limited set of aspects about proxy model development, and some of them report results which are contradictory to the conclusions of previous work. A good example is the proxy model development process and its execution, where there were no reductions in computational time because it depended on the problem to be more efficient than the application with a commercial simulator. To summarise, this SLR in this field is a valuable source for researchers and interested parties in the development of the proxy model.

### 5.4 Why use the proxy model?

Numerical reservoir simulators are used at various stages of field development and management phases in the oil and gas industry. Petroleum reservoir engineers evaluate the fluid behaviour and drainage patterns during the production using reservoir simulation models. This procedure is related to the three main areas of field development and management phases, illustrated in Fig. 6.

Reservoir simulation is an essential tool for reservoir studies because it permits the representation of reality (real petroleum field) through a physical model which can be used to describe petroleum production under various operating conditions. Depending on the complexity (size and representativeness) of the model, the reservoir simulation process demands high computational time and resources. The high-level heterogeneity of reservoirs and fluid-type injected to increase petroleum recovery factor often requires high-fidelity models to represent the reality in numerical simulation. Decision analysis related to the management of petroleum fields with high-fidelity models is time-consuming, mainly in probabilistic approaches to

cover all possible solutions. Additionally, the decision involves a risk analysis that accounts for several uncertainties, mostly in the reservoir and fluid parameters, economic model, operational availability, and high computational cost.

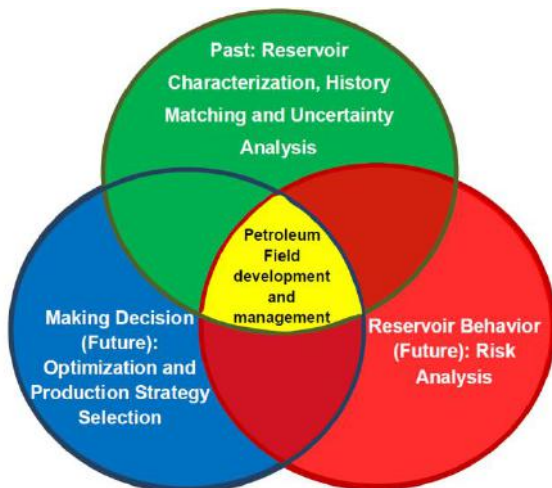


Fig. 6 The three main areas of petroleum reservoir studies where it is possible to apply proxy model techniques.

The authors [1] portrayed the complexity of the process based on 12-step closed-loop reservoir development and management. Several research fields that suggested the investigation of appropriated architectures and methodologies were used as proxies to accelerate some parts of the process. The authors [127] referred to the proxy model as “metamodels”; in other words, it is a “model of a model”. “Model emulation” is another term referring to surrogate modelling (proxy model) [128]. The authors [129, 130] mentioned that the term response surface surrogate in the literature is referring to the metamodel. This way, the proxy model can be defined, such as an approximation of a response function built using data fitting of limited simulation results [131].

Moreover, a metamodel is a relatively simple model used to mimic the reservoir simulator output, reproducing the simulation’s input-output relationships. The quality of the proxy model generated will depend on the mathematical approach, and the input used to build it. There are many motivations to create the proxy model, such as [132-134]:

- Better use of the available, typically limited, computational budget;
- Low-resolution models for simple analysis (predict future petroleum production);

- The models (input and output) are often large and complex;
- Computational demands result in high computer time for obtaining results from such complex models, especially in probabilistic settings;
- Unreasonably high computer times could prevent decision-maker from exploring the design space, resulting in underperforming results.

The main obstacle of the reservoir numerical simulator is the extensive use of the most sophisticated techniques, and the high number of model runs required. On the other hand, the proxy models tend to be fast.

The most used proxy models in the oil industry highlighted in the SLR were: kriging model (KG); artificial neural network (ANN); response surface methodology (RSM), fuzzy logic (FL), Kalman filter (KF), Experimental Design (ED), and Bayesian emulators (EM) and other models such as genetic algorithm (GA), Karhunen–Loève expansion (KL), polynomial chaos expansion (PCE), support vector machine (SVM) and deep learning (DL).

#### 5.4.1 Kriging Model

Kriging (KG) is a geostatistical technique for estimating properties at locations that do not have measured data [55]. In other words, KG is the geostatistical method of predicting values at unsampled points [135], which is a form of multi-dimensional interpolation very commonly used to build the proxy model in petroleum reservoir studies. It uses a variogram model (a measure of spatial correlation) to infer the weights given to each data point.

It is worth mentioning that KG is similar to other interpolation methods, such as radial basis function (RBF) and spline. Besides, it is a combination of a polynomial model, which is a global function over the entire input space, and a localized deviation model based on spatial correlation of samples [133].

According to [135], the main goal of the KG is to predict the values of stationary covariance at the unsampled point concerning the mean squared error. The covariance function is not commonly known and needs to be estimated. There are some types of KG: ordinary kriging, simple kriging, universal kriging, and the co-kriging, [135] presented the details and their mathematical derivation. The authors [55, 135] give more information on this technique.

#### 5.4.2 Artificial Neural Network

Artificial neural networks (ANNs) are structures inspired by biological nervous systems, which can deal with different complex problems. In other words, ANNs are computational models developed on the principle of the biological nervous system [64]. According to [65], it is a virtual intelligence or machine learning technique which is useful for pattern recognition and prediction of a complicated non-linear relationship between input and output.

ANNs can assimilate highly complex relationships between several variables that are presented to the network and learn the characteristics of the dependency between input and output [62, 63]. ANNs are classified in supervised and unsupervised learning. Unsupervised learning is used to classify a set of data into a specific number of features. In contrast, supervised learning classifies patterns and makes decisions based on the patterns of inputs and outputs learned.

The use of ANNs has been increasing in the oil and gas industry over the past decades to solve many complex and highly non-linear problems [136] and uncertain relationships between the input and output for given dataset [68]. According to [61], the results of some applications of ANNs in several research fields suggest the investigation of appropriated architectures for reservoir simulator. They have been successfully applied in several research fields of petroleum engineering to solve various problems, for example, reservoir characterisation, forecasting, risk analysis, history matching, uncertainty analysis, optimisation, production strategy selection, among others. The authors [3, 69] present more application of ANNs in the oil and gas industry.

The difficulty in the application of ANNs as a reservoir simulator proxy is for them to be fully trained, which requires a large number of reservoir simulation runs [61]. Otherwise, ANNs have the benefit over other conventional techniques, such as response surface and reduced models, to perform complex and highly non-linear inputs and outputs accurately and rapidly [69]. According to [137], ANNs offer some advantages, including their capacity of inferring highly complex, nonlinear, and possibly uncertain relationships between system variables, requiring practically zero prior knowledge regarding the unknown function.

#### 5.4.3 Response Surface Methodology

Response Surface Methodology (RSM) is an application of statistical and mathematical techniques useful for developing and optimizing models and parameters [18]. The authors [139] defined RSM as a

combination of statistical methods to build an empirical model for the objective function used in the process.

The authors [69] highlighted various studies which used RSM to calculate the porosity and permeability distribution in a heterogeneous and multiphase reservoir. Also, to replicate the results of a full field simulation model based on time complexity, and to analyse of the uncertainty of coalbed methane production to optimise the performance of a reservoir; among other studies. According to [40], the goal of the experimental design and RSM is to build response surfaces of specific objective functions that genuinely represent the response. For more application using RSM in the oil and gas industry, see [31].

#### 5.4.4 Fuzzy Logic

Fuzzy logic (FL) is a superset of conventional Boolean logic that has been extended to handle the concept of partial values between true and false [139, 140]. In other words, FL is logic or probabilistic form, which deals with reasoning that is approximate rather than exact. It is built with fuzzifier, the inference mechanism, the rules, and the defuzzifier.

In the petroleum industry, there are many different studies with the application of FL, for example, [141] in dealing with the uncertainty of a number introduced a fuzzy analytic hierarchy process. This process describes a relationship between an uncertain quantity and a function which ranges from 0 to 1. The authors [49] present more studies concerning FL.

#### 5.4.5 Kalman Filter

According to [142], the Kalman filter (KF) can be viewed, such as a Bayesian estimator that approximates conditional probability densities of the time-dependent vector. KF is optimal for linear problems for assimilating measurements to update the estimate of variables continuously. Additionally, it is most appropriate when a short number of variables characterizes the issues and when the variables to be estimated are linearly related to the observations [21, 23]. According to [23], this case is not applied to spatiotemporal reservoir problems because the number of model parameters is typically very high, and the relation between the reservoir model and the production observations, represented by a fluid-flow simulator, is highly nonlinear. It is essential to highlight that most data assimilation problems in petroleum reservoir engineering are highly nonlinear and are characterised by many variables.

Several extensions of the KF techniques have been suggested, such as ensemble Kalman filter (EnKF),



developed by [143], and documented by [144-146]. In reservoir engineering literature, EnKF has been primarily used to estimate or stochastically simulate grid block permeabilities and porosities [147]. Therefore, it can be conceptually extended to include other parameters [22].

According to [26], EnKF performs the initial sampling, forecasting, and assimilation steps for automatic history matching in the petroleum industry. EnKF has emerged as an attractive option for reservoir history matching problems because it is simple to implement and can be computationally efficient [27-30; 147]. It can also improve the accuracy and reduce the corresponding predictive uncertainty by accounting for observations [9].

The use of the EnKF is a promising approach for data assimilation and assessment of uncertainties during reservoir characterization and performance forecasting [25]. Many studies using EnKF in petroleum engineering can be seen in [22, 23, 26].

#### 5.4.6 Experimental Design

Experimental Design (ED) can be used to generate a reliable response surface which covers the entire range of uncertain parameters [3]. In other words, according to [148], ED presents a method that investigates the effects of multiple variables on output or response, simultaneously.

The experiment of [39], with the ED application, involved many simulations and are made changes on the input variable. The authors [5] mentioned that, in an experiment, one or more variables could be changed to quantify the effect of inputs on outputs (response variables). ED is used to avoid the time-consuming process to capture all changes with the minimum number of simulator runs [31, 38]. The authors [34, 38, 41] show many studies in petroleum engineering which applied the ED methodology.

#### 5.4.7 Bayesian Emulators

The authors [10] inform that an emulator is usually composed of a predictor (a statistical approximation of the unknown function), and also by predictor uncertainty quantification. In other words, an emulator uses reservoir properties as input in a statistical model constructed from simulator outputs. The emulator response is faster, but there is still a need to establish the issues with uncertainty in the inputs and outputs.

The number of methodologies using Bayesian emulators is increasing [11-13, 62, 149-154]. But, there are still some obstacles in the implementation, especially in production strategy selection stages, as follows:

- The high computational costs in the quantification of probabilistic problems;

- Effective ways to parameterize the geostatistical realization uncertainties (porosity and permeability distribution);
- Analysis of measurement errors of various classes of uncertainties;
- Assessment of model discrepancy for uncertainty quantification;
- Practical techniques for the decision-making process.

This way, the development of emulators requires careful consideration of various factors, such as optimization process, uncertainty quantification and computational time. The initial knowledge of factor effects during the emulator's construction is fundamental to obtain a useful emulator.

It is worth highlighting that there are many uncertainties associated with the generation of geostatistical realizations [8, 155]. These are combined with realizations from the reservoir, technical and/or economic models to compose the different reservoir model scenarios [156]. These scenarios are then used to make decisions without fully accounting for uncertainties and risks.

#### 5.4.8 Other Proxy Models

This section includes a summary of other proxy models found in the SLR.

- Genetic algorithm (GA) – They are stochastic search and optimization heuristics methods from classical evolution theory [157]. Moreover, they require only objective function evaluations to find optimised points, even though the derivative information is not available [48]. Therefore, their extensive application in different fields is proof that they can be applied to various engineering problems [48, 60].
- Karhunen–Loeve expansion (KL) - It is a promising approach for representing random fields from a covariance matrix. It is a linear relation that decorrelates the random field while preserving the two-point statistics of the area [7]. The covariance function may describe the correlation structure of the random field. KL is an optimal technique for parameterization [158] because it approximates the original random area accurately and with a minimum number of inputs [7]. The authors [159] present more details about KL.
- Polynomial Chaos Expansion (PCE) – Wiener (1938) introduced this technique. According to [59], PCE obtained notable of popularity for the uncertainty quantification of dynamic systems. It is worth mentioning that [159] were pioneers in the use of uncertainty quantification. Nowadays, PCE is applied to various problems and studies in petroleum engineering. The



authors [159] used PCE to quantify uncertainty for efficient closed-loop production optimisation. While the authors [59] used PCE as a proxy substitute for the full reservoir simulator proxy when applied to the Markov chain Monte Carlo method and the authors [7] used PCE to predict the production parameters of steam-assisted gravity drainage (SAGD) reservoir. It is worth noting that PCEs have a significant advantage over other proxy models, because of their convergence in probability and distribution to the output random variable of interest [7, 59].

- Support Vector Machine (SVM) – It is a part of machine learning (artificial intelligence – AI), a supervised learning technique, being widely applied in classification and regression analysis. According to [31], AI is an application in the oil and gas industry which has enormous potential to explore the knowledge regarding reservoir characterization, PVT properties, well placement, etc. The authors [8, 161-163] presented an application in the petroleum engineering.
- Deep learning (DL) – We did not identify this technique in the articles analyzed for the development of SLR, but some authors mentioned future work utilizing DL in petroleum engineering. The authors [164, 165] applied DL to petroleum well data.

## VI. CONCLUSION

This research enables us not only to know about state-of-the-art proxy modelling but also serves to identify the primary contexts in which to apply it. Besides, it provides us with insight into the criteria used when facing the need to decide on the method based on a context to perform this task. In this SLR, we identified the three main area applications related to the petroleum engineering area: past, future (decision-making) and future (reservoir behaviour, production forecast). These area applications are based in 6 topics (three past, two decision-making and one reservoir behaviour and production forecast): uncertainty analysis, history matching, reservoir characterisation, optimisation, production strategy selection and risk analysis.

Depending on the complexity of the model, the use of reservoir simulator is more efficient than a proxy because of the high computational time and human resources. A total of 64.96% of the 117 publications selected, the authors mentioned that the computational time reduction is essential for the development of the proxy model development. When working on proxy modelling, this becomes even more complex due to high-heterogeneous reservoirs and high-dimensionality problems, especially in

maintaining the geological consistency, which is the main focus of reservoir modelling. The dimensionality reduction is a complex problem and involves thousands of reservoir simulation runs, representing an obstacle for practical applications when we did not define the adequate method and number of dimension.

This SLR has various limitations, mainly in the petroleum engineering area, because it is not a developed research method. Another limitation is the inclusion and exclusion criteria constructed and used in our research. This research relies on certain types of publications in reviewing academic literature. We did not include in the development of the SLR, the scientific articles published as books, technical reports, work in progress, and publications without bibliographical information or unpublished research that were not in the seven databases. Therefore, this research may be missing relevant studies published in other digital libraries, or those did not appear in the search results due to the search string. Due to criteria, we were in line with the exclusion criteria of this study, and with all requirements established systematically, so as not to pose risks for validating the results.

The primary purpose of this SLR was to ascertain existing decision-making and criteria for the comparison and selection of methods for proxy model development in future research. The results may also be useful for researchers as it can help them to analyse the existing publications in the different methods utilized in the metamodel development, identifying gaps to perform further research. Additionally, from the SLR, scientific methods are straightforward and reproducible because their proposed methodology enables an accurate survey and a scientific development of the state-of-the-art in the specific problems of research. For this reason, we could achieve future work about metamodel with the integration of fast methods and reservoir numerical simulator runs. The integration can improve and accelerate results within acceptance criteria and accuracy in the decision-making process related to reservoir management and development, which are necessary for the uncertainty quantification process in the petroleum field.

## ACKNOWLEDGEMENTS

This work was conducted with the support of Energi Simulation and in association with the ongoing Project registered as "BG-32 – Análise de Risco para o Desenvolvimento e Gerenciamento de Campos de Petróleo e Potencial uso de Emuladores" (UNICAMP/Shell Brazil/ANP) funded by Shell Brazil, under the ANP R&D levy as "Compromisso de Investimentos com Pesquisa e

Desenvolvimento". The authors also thank UNISIM, DE-FEM-UNICAMP, CEPETRO and FUNCAMP for supporting this work, and CNPq for (Research and Development National Council) for the financial support for the development of the project under grant process number 200020/2019-6.

## REFERENCES

- [1] Schiozer, D. J., Santos, A. A. S, Drumond, P. S (2015) Integrated model based decision analysis in twelve steps applied to petroleum fields development and management. SPE-174370-MS, EUROPEC, Madrid. <https://doi.org/10.2118/174370-MS>.
- [2] Pinto, M. A. S, Ghasemi, M., Gildin, E., Schiozer, D. J (2015). Hybrid optimization for closed-loop reservoir management. SPE-173278-MS, Reservoir Simulation Symposium, Texas. <https://doi.org/10.2118/173278-MS>.
- [3] Panjalizadeh, H., Alizadeh, N., Mashhadi, H (2014). Uncertainty assessment and risk analysis of steam flooding by proxy models, a case study. International Journal of Oil Gas and Coal Technology, 7(1), 29-51. <https://doi.org/10.1504/IJOGCT.2014.057795>.
- [4] Yeten, B., Castellini, A., Guyaguler, B., Chen, W. H (2005). A comparison study on experimental design and response surface methodologies. SPE 93347, Reservoir Simulation Symposium, Texas. <https://doi.org/10.2118/93347-MS>.
- [5] Norouzi, M., Panjalizadeh, H., Rashidi, F., Mahdiani, M. R (2017). DPR polymer gel treatment in oil reservoirs: A workflow for treatment optimization using static proxy models. Journal of Petroleum Science and Engineering, 153, 97-110. <https://doi.org/10.1016/j.petrol.2017.03.018>.
- [6] Risso, F. V. A., Risso, F. F., Schiozer, D. J (2008). Risk assessment of oil fields using proxy models: A case study. Journal of Canadian Petroleum Technology, 47(8), 9-14. <https://doi.org/10.2118/08-08-09-TN>.
- [7] Jain, T., Patel, R. G., Trivedi, J (2017). Application of polynomial chaos theory as an accurate and computationally efficient proxy model for heterogeneous steam-assisted gravity drainage reservoirs. Energy Science and Engineering, 5(5), 270-289. <https://doi.org/10.1002/ese3.177>.
- [8] Da Silva, L. M., Avansi, G. D., Schiozer, D. J. (2020). Support Vector Regression for Petroleum Reservoir Production Forecast Considering Geostatistical Realizations. SPE Reservoir Evaluation & Engineering, 24, 1-15. <https://doi.org/10.2118/203828-PA>.
- [9] Xue, L., Dai, C., Wang, L (2017). Development of a general package for resolution of uncertainty-related issues in reservoir engineering. Energies, 10, 1-16. <https://doi.org/10.3390/en10020197>.
- [10] Busby, D., Feraille, M (2008). Adaptive design of experiments for calibration of complex simulators - An application to uncertainty quantification of a mature oil field. Journal of Physics: Conference Series, 135, 1-8. <https://doi.org/10.1088/1742-6596/135/1/012026>.
- [11] Slotte, P. A., Smørgrav, E (2008). Response surface methodology approach for history matching and uncertainty assessment of reservoir simulation models. 70th European Association of Geoscientists and Engineers Conference and Exhibition 2008: Leveraging Technology. Incorporating SPE EUROPEC, 3, 1408-1417. <https://doi.org/10.2118/113390-MS>.
- [12] Stordal, A. S., Nævdal, G (2017). A modified randomized maximum likelihood for improved Bayesian history matching. Computational Geosciences, 22, 29-41. <https://doi.org/10.1007/s10596-017-9664-x>.
- [13] Dang, C., Nghiem, L., Nguyen, N., Zhangxin, C., Yang, C., Nguyen, Q (2017). A framework for assisted history matching and robust optimization of low salinity waterflooding under geological uncertainties. Journal of Petroleum Science and Engineering, 152, 330-352. <https://doi.org/10.1016/j.petrol.2017.03.009>.
- [14] Ding, Y., Renard, G., Herzhaft, B (2008). Quantification of uncertainties for drilling-induced formation damage. SPE Journal, 221-231. <https://doi.org/10.2118/100959-PA>.
- [15] Feraille, M., Busby, D (2009). Uncertainty management on a reservoir workflow. Society of Petroleum Engineers - International Petroleum Technology Conference, 4, 2507-2522. <https://doi.org/10.2523/IPTC-13768-MS>.
- [16] Metla, N., Delbos, F., Da Veiga, S., Sinoquet, D (2010). Constrained nonlinear optimization for extreme scenario evaluation in reservoir characterization. ECMOR - 12th European Conference on the Mathematics of Oil Recovery. <https://hal-ifp.archives-ouvertes.fr/hal-02284357>.
- [17] Azad, A., Chalaturnyk, R. J (2013). Application of analytical proxy models in reservoir estimation for the SAGD process: UTF-Project case study. Journal of Canadian Petroleum Technology, 52(3), 219-232. <https://doi.org/10.2118/165576-PA>.
- [18] Dai, Z., Middleton, R., Viswanathan, H., Fessenden-Rahn, J., Bauman, J., Pawar, R., Lee, S-Y., McPherson, B (2014). An integrated framework for optimizing CO2 sequestration and enhanced oil recovery. Environmental Science & Technology Letters, 1, 49-54. <https://doi.org/10.1021/ez4001033>.
- [19] Dehdari, V., Deutsch, C. V (2015). Optimizing well trajectories in steam-assisted-gravity-drainage reservoir development. SPE Reservoir Evaluation & Engineering, 18(1), 53-68. <https://doi.org/10.2118/174078-PA>.
- [20] Valladão, D. M., Torrado, R. R., Flach, B., Embid, S (2013). On the stochastic response surface methodology for the determination of the development plan of an oil & gas field. Society of Petroleum Engineers – SPE: Intelligent

- Energy International, 534-545. <https://doi.org/10.2118/167446-MS>.
- [21] Gu, Y., Oliver, D. S (2007). An iterative ensemble Kalman Filter for multiphase fluid flow data assimilation. SPE Journal, 438-446. <https://doi.org/10.2118/108438-PA>.
- [22] Thulin, K., Li, G., Aanonsen, S.I., Reynolds, A.C (2007). Estimation of initial fluid contacts by assimilation of production data with EnKF. Proceedings - SPE Annual Technical Conference and Exhibition, 3, 1655-1669 <https://doi.org/10.2118/109975-MS>.
- [23] Lødøen, O. P., Omre, H (2008). Scale-Corrected Ensemble Kalman Filtering Applied to Production-History Conditioning. Reservoir Evaluation. SPE-111374-PA.13(2), 177-194. <https://doi.org/10.2118/111374-PA>.
- [24] Chen, Y., Oliver, D. S., Zhang, D (2009). Efficient ensemble-based closed-loop production optimization. SPE Journal, 14(4), 634-645. <https://doi.org/10.2118/112873-PA>.
- [25] Arroyo-Negrete, E., Devegowda, D., Datta-Gupta, A., Choe, J (2008). Streamline-assisted ensemble Kalman Filter for rapid and continuous reservoir model updating. SPE Journal, 1046-1060. <https://doi.org/10.2118/104255-PA>.
- [26] Liang, B., Sepehrmoori, K., Delshad, M (2009). A weighted ensemble Kalman Filter for automatic history matching. Petroleum Science and Technology, 27(10), 1062-1091. <https://doi.org/10.1080/10916460802455939>.
- [27] Jafarpour, B., McLaughlin, D. B (2009). Estimating channelized-reservoir permeabilities with the ensemble Kalman Filter: The importance of ensemble design. SPE Journal, 14(2), 374-388. <https://doi.org/10.2118/108941-PA>.
- [28] Li, G., Han, M., Banerjee, R., Reynolds, A. C (2009) Integration of well test pressure data into heterogeneous geological reservoir models. Proceedings - SPE Annual Technical Conference and Exhibition, 2, 889-902. <https://doi.org/10.2118/124055-MS>.
- [29] Emerick, A. A., Reynolds, A. C (2012). Combining the ensemble Kalman Filter with Markov Chain Monte Carlo for improved history matching and uncertainty characterization. SPE Journal, 17(2), p. 418-440. <https://doi.org/10.2118/141336-MS>.
- [30] Jahanbakhsh, A., Elsheikh, A., Sohrabi, M (2016). Application of ensemble smoother and multiple data assimilation for estimating relative permeability from core flood experiments. 15th European Conference on the Mathematics of Oil Recovery, ECMOR. <https://doi.org/10.3997/2214-4609.201601816>.
- [31] Panja, P., Pathak, M., Velasco, R., Deo, M (2016). Least square support vector machine: An emerging tool for data analysis. Society of Petroleum Engineers - SPE Low Perm Symposium. <https://doi.org/10.2118/180202-MS>.
- [32] Tang, H., Wang, F (2007). Using production data to mitigate reservoir connectivity uncertainty. International Petroleum Technology Conference, IPTC, 1, 97-107. <https://doi.org/10.2118/2007-026-EA>.
- [33] Avansi, G. D., Schiozer, D. J., Suslick, S. B., Risso, F. V. A (2009) Assisted Procedures for Definition of Production Strategy and Economic Evaluation Using Proxy Models. SPE Europec/EAGE Annual Conference and Exhibition, Amsterdã, Holanda. SPE 122298.
- [34] Gupta, K., Collinson, R., Smith, G. C., Ryan, S., Louis, J (2008). History matching of field production using design of experiments. SPE Asia Pacific Oil and Gas Conference and Exhibition - "Gas Now: Delivering on Expectations, 2, 862-868. <https://doi.org/10.2118/115685-MS>.
- [35] Prada, J. W. V., Cunha, L. B (2008). Assessment of optimal operating conditions in a SAGD project by the design of experiments and response surface methodology. Petroleum Science and Technology, 26(17), 2095-2107. <https://doi.org/10.1080/10916460701429399>.
- [36] Prada, J. W. V., Cunha, L. B (2008). Prediction of SAGD performance using response surface correlations developed by experimental design techniques. Journal of Canadian Petroleum Technology, 47, 58-64. <https://doi.org/10.2118/08-09-58>.
- [37] Schaaf, T., Coureaud, B., Labat, N., Busby, D (2009). Using experimental designs assisted history-matching tools, and Bayesian framework to get probabilistic gas-storage pressure forecasts. SPE Reservoir Evaluation & Engineering, 12(5), 724-736. <https://doi.org/10.2118/113498-MS>.
- [38] Moeinikia, F., Alizadeh, N (2012). Experimental design in reservoir simulation: an integrated solution for uncertainty analysis, a case study. Journal of Petroleum Exploration and Production Technology, 2(2), 75-83. <https://doi.org/10.1007/s13202-012-0023-0>.
- [39] Ajibola, J. T., Orodu, O. D., Onyeukwu, C. A (2013). Sidetrack/recompletion time evaluation by proxy model. Society of Petroleum Engineers - 37th Nigeria Annual Int. Conf. and Exhibition (NAICE). To Grow Africa's Oil and Gas Production: Required Policy, Funding, Technol., Techniques and Capabilities, 2, 1105-1113. <https://doi.org/10.2118/167593-MS>.
- [40] Al-Shalabi, E. W., Sepehrmoori, K., Delshad, M (2014). Optimization of the low salinity water injection process in carbonate reservoirs. Society of Petroleum Engineers - International Petroleum Technology Conference. Innovation and Collaboration: Keys to Affordable Energy, 2, 1082-1109. <https://doi.org/10.2523/IPTC-17821-MS>.
- [41] Bhark, E., Dehghani, C (2014). Assisted history matching benchmarking: Design of experiments-based techniques. Proceedings - SPE Annual Technical Conference and Exhibition, 2, 1454-1488. <https://doi.org/10.2118/170690-MS>.

- [42] Ligeró, E. L., Alves Risso, F. V., Schiozer, D. J (2007). Comparison of methodologies to evaluate the risk of petroleum fields. Proceedings of the SPE Latin American and Caribbean Petroleum Engineering Conference, 2, 1016-1023. <https://doi.org/10.2118/107736-MS>.
- [43] Bauman, J. H., Deo, M. D (2011). Parameter space reduction and sensitivity analysis in complex thermal subsurface production processes. Energy & Fuels, 25, 251-259. <https://doi.org/10.1021/ef101225g>.
- [44] Maučec, M., Douma, S., Hohl, D., Leguijt, J., Jimenez, E.A., Datta-Gupta, A (2007). Streamline-based history matching and uncertainty: Markov-Chain Monte Carlo study of an offshore turbidite oil field. Proceedings - SPE Annual Technical Conference and Exhibition, 3, 1506-1521. <https://doi.org/10.2118/109943-MS>.
- [45] Maučec, M., Cullick, S., Shi, G (2011). Geology-guided quantification of production-forecast uncertainty in dynamic model inversion. Proceedings - SPE Annual Technical Conference and Exhibition, 3, 2496-2512. <https://doi.org/10.2118/146748-MS>.
- [46] Maučec, M., Cullick, S., Shi, G (2011). Quantitative uncertainty estimation and dynamic model updating for improved oil recovery. Society of Petroleum Engineers - SPE Enhanced Oil Recovery Conference, 1, 576-590. <https://doi.org/10.2118/144092-MS>.
- [47] Scheidt, C., Caers, J (2009). Uncertainty quantification in reservoir performance using distances and Kernel Methods-application to a west Africa Deepwater turbidite reservoir. SPE Journal, 680-692. <https://doi.org/10.2118/118740-PA>.
- [48] Bello, O., Awofodu, D. D., Oppelt, D. D. J., Ganzer, L., Holzmann, J (2017). Selective perforation & design of multi-pattern infill wells in field development planning & optimization under geological uncertainty. Proceedings of the Annual Offshore Technology Conference, 6, 4265-4281. <https://doi.org/10.4043/27553-MS>
- [49] Qin, T., Chen, Z., Zhang, K., Han, J., Wu, J. K (2016). Redevelopment of the Pembina Cardium field by CO<sub>2</sub>-EOR using existing wells. Society of Petroleum Engineers - SPE Europec Featured at 78th EAGE Conference and Exhibition. <https://doi.org/10.2118/180178-MS>.
- [50] Alvi, A., Tilke, P., Bogush, A., Kolupaev, A., Jilani, S.Z., Banerjee, R., Bolanos, N., Wu, J (2012). Delivering optimal brownfield development strategies: Use of multi-constrained optimization combined with fast semi-analytical and numerical simulators. Society of Petroleum Engineers - Kuwait International Petroleum Conference and Exhibition: People and Innovative Technologies to Unleash Challenging Hydrocarbon Resources, 2, 617-626. <https://doi.org/10.2118/163337-MS>.
- [51] Sabatino, R., Viviani, E., Della Rossa, E., Sala, C., Maffioli, A (2014). Structural uncertainty integration within reservoir risk analysis and history matching. Proceedings - SPE Annual Technical Conference and Exhibition, 4, 2490-2500. <https://doi.org/10.2118/170761-MS>.
- [52] Busby, D., Feraille, M (2008). Dynamic data assimilation by MCMC and sequential design of experiments. ECMOR - 11th European Conference on the Mathematics of Oil Recovery. <https://doi.org/10.3997/2214-4609.20146416>.
- [53] Crumpton, P. I., Habiballah, W. A., Wardell-Yerburgh, P.G., Nasser, K. A., Faleh, A. A (2011). Multilateral-complex well optimization. SPE Reservoir Simulation Symposium, 1, 80-91. <https://doi.org/10.2118/140882-MS>.
- [54] Nguyen, N. T. B., Chen, Z., Dang, C. T. Q., Nghiem, L. X., Yang, C., Bourgoult, G., Li, H (2015). Integrated modeling for assisted history matching and robust optimization in mature reservoirs. SPE/IATMI Asia Pacific Oil and Gas Conference and Exhibition, APOGCE. <https://doi.org/10.2118/176290-MS>.
- [55] Dang, C., Nghiem, L., Nguyen, N., Zhangxin, C., Nguyen, Q (2016). Evaluation of CO<sub>2</sub> low salinity water-alternating-gas for enhanced oil recovery. Journal of Natural Gas Science and Engineering, 35, 237-258. <https://doi.org/10.1016/j.jngse.2016.08.018>.
- [56] Li, H., Zhang, D (2009). Efficient and accurate quantification of uncertainty for multiphase flow with the probabilistic collocation method. SPE Journal, 14(4), 665-679. <https://doi.org/10.2118/114802-PA>.
- [57] Goldstein, M (2012). Bayes linear analysis for complex physical systems modeled by computer simulators. IFIP Advances in Information and Communication Technology - AICT, 377, 78-94. [https://doi.org/10.1007/978-3-642-32677-6\\_6](https://doi.org/10.1007/978-3-642-32677-6_6).
- [58] Vanegas, P. J. W., Clayton, P., Deutsch, C. V., Cunha, L. B (2008). Uncertainty assessment of SAGD performance using a proxy model based on Butler's theory. Proceedings - SPE Annual Technical Conference and Exhibition, 3, 1709-1729. <https://doi.org/10.2118/115662-MS>.
- [59] Bazargan, H., Christie, M., Tchelepi, H (2013). Efficient Markov Chain Monte Carlo sampling using polynomial chaos expansion. SPE Reservoir Simulation Symposium, 2, 1183-1203. <https://doi.org/10.2118/163663-MS>.
- [60] Ali-Ahmad, M., Zendehboudi, S., James, L. A., Elkamel, A., Dusseault, M., Chatzis, I., Lohi, A (2014). New tools to determine bubble point pressure of crude oils: experimental and modeling study. Journal of Petroleum Science and Engineering, 123, 207-216. <https://doi.org/10.1016/j.petrol.2014.08.018>.
- [61] Silva, P. C., Maschio, C., Schiozer, D. J (2007). Use of neuro-simulation techniques as proxies to reservoir simulator: Application in production history matching. Journal of Petroleum Science and Engineering, 57, 273-280. <https://doi.org/10.1016/j.petrol.2006.10.012>.
- [62] Maschio, C., Schiozer, D. J (2014). Bayesian history matching using artificial neural network and Markov Chain Monte Carlo. Journal of Petroleum Science and



- Engineering, 123, 62-71. <https://doi.org/10.1016/j.petrol.2014.05.016>.
- [63] Costa, L. A. N., Maschio, C., Schiozer, D. J (2014). Application of artificial neural networks in a history matching process. *Journal of Petroleum Science and Engineering*, 123, 30-45. <https://doi.org/10.1016/j.petrol.2014.06.004>.
- [64] Srinivasan, K., Ertekin, T (2008). Development and testing of an expert system for coalbed methane reservoirs using artificial neural networks. *Society of Petroleum Engineers - SPE Eastern Regional/AAPG Eastern Section Joint Meeting*, 597-606. <https://doi.org/10.2118/119935-MS>.
- [65] Amirian, E., Leung, J. Y., Zanon, S., Dzurman, P (2013). Data-driven modelling approach for recovery performance prediction in SAGD operations. *Society of Petroleum Engineers - SPE Heavy Oil Conference Canada*, 3, 2206-2222. <https://doi.org/10.2118/165557-MS>.
- [66] Memon, P. Q., Yong, S.-P., Pao, W., Sean, P. J (2014). Surrogate reservoir modeling-prediction of bottom-hole flowing pressure using a radial basis neural network. *Science and Information Conference*, 499-504. <https://doi.org/10.1109/SAI.2014.6918234>.
- [67] Ma, Z., Leung, J. Y., Zanon, S., Dzurman, P (2015). Practical implementation of knowledge-based approaches for steam-assisted gravity drainage production analysis. *Expert Systems with Applications*, 42(21), 7326-7343. <https://doi.org/10.1016/j.eswa.2015.05.047>.
- [68] Ma, Z., Liu, Y., Leung, J. Y., Zanon, S (2015). Practical data mining and artificial neural network modelling for SAGD production analysis. *SPE Canada Heavy Oil Technical Conference*, 1178-1202. <https://doi.org/10.2118/174460-MS>.
- [69] Memon, P. Q., Yong, S.-P., Pao, W., Pau, J. S (2015). Dynamic Well Bottom-Hole Flowing Pressure Prediction Based on Radial Basis Neural Network. *Intelligent Systems in Science and Information*, 279-292. [https://doi.org/10.1007/978-3-319-14654-6\\_17](https://doi.org/10.1007/978-3-319-14654-6_17).
- [70] Kitchenham, B., Charters, S (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. *EBSE Technical Report*, Keele, UK. <https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.471>.
- [71] Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., Linkman, S (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51, 7-15. <https://doi.org/10.1016/j.infsof.2008.09.009>.
- [72] Ferreira, L. M. Alves-Souza, S., Filgueiras, L. Multidimensional Modelling in NoSQL Database: A Systematic Review. *IEEE Access*, 16p. In review.
- [73] EIA – Independent Statistics & Analysis (U.S. Energy Information Administration). Petroleum & other liquids. Data: Spot Prices (Crude Oil in Dollars per Barrel, Products in Dollars per Gallon). Available in <https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PE&t=s=RBRTE&f=M>
- [74] Camacho Velazquez, R., Fuentes-Cruz, G., Vasquez-Cruz, M. A (2008). Decline-Curve Analysis of Fractured Reservoirs With Fractal Geometry. *SPE*, 606-619. <https://doi.org/10.2118/104009-PA>.
- [75] Marques, J. B. D., Trevisan, O. V., Marques, L. M., Ravagnani, A. T. G (2014). Stochastic processes and copula model applied in the economic evaluation for Brazilian oil fields projects. *SPE Hydrocarbon Economics and Evaluation Symposium*, 379-399. <https://doi.org/10.2118/169857-MS>.
- [76] Valiullin, R., Ramazanov, A., Sadretdinov, A., Sharafutdinov, R., Shako, V., Sidorova, M., Kryuchatov, D (2014). Field study of temperature simulator applications for quantitative interpretation of transient thermal logging in a multipay well. *Society of Petroleum Engineers - SPE Russian Oil and Gas Exploration and Production Technical Conference and Exhibition - Sustaining and Optimising Production: Challenging the Limits with Technology*, 1, 827-852. <https://doi.org/10.2118/171233-RU>.
- [77] Nandanwar, M. S., Anderson, B. J., Ajayi, T., Collett, T. S., Zyrianova, M. V (2016). Evaluation of gas production potential from gas hydrate deposits in National Petroleum Reserve Alaska using numerical simulations. *Journal of Natural Gas Science and Engineering*, 36, 760-772. <https://doi.org/10.1016/j.jngse.2016.11.021>.
- [78] Khajooie, S., Roghanian, R., Shahrabadi, A., Tavakkoli, M (2012). The Validation of Methods to Calculate Gas-Oil Relative Permeability Using Capillary Pressure Data from South Pars Reservoir. *Petroleum Science and Technology*, 30(23), 2424-2434. <https://doi.org/10.1080/10916466.2010.519750>.
- [79] Frooqnia, A., Torres-Verdin, C., Sepehrnoori, K., A-Pour, R (2017). Inference of near-borehole permeability and water saturation from time-lapse oil-water production logs. *Journal of Petroleum Science And Engineering*, 152, 116-135. <https://doi.org/10.1016/j.petrol.2017.03.005>
- [80] Sancho, V., Berrios, V., Gulino, F., Oquendo, O (2007). A field development strategy: When a simple stochastic model surpasses numerical simulation. *Proceedings of the SPE Latin American and Caribbean Petroleum Engineering Conference*, 3, 1490-1496. <https://doi.org/10.2118/108061-MS>.
- [81] Batyrshin, I (2011). Uncertainties with Memory in Construction of Strict Monotonic T-Norms and T-Conorms for Finite Ordinal Scales: Basic Definitions and Applications. *Applied and Computational Mathematics*, 10(3), 498-513.
- [82] Chassagne, R. L., Hammond, P. S (2012). Simulation of Drilling Fluid Filtrate Invasion Near an Observation Well. *SPE Journal*, 17(4), 1047-1055. <https://doi.org/10.2118/154014-PA>.



- [83] Pirrone, M., Cossa, A., Galli, M. T (2014). Probabilistic high-resolution permeability estimation integrating dielectric dispersion log measurements with mud filtrate invasion modelling. Society of Petroleum Engineers - IPTC - Innovation and Collaboration: Keys to Affordable Energy, 3, 2267-2279. <https://doi.org/10.2523/17967-MS>.
- [84] Khrulenko, A., Zolotukhin, A. B (2011). Approach for full-field scale smart well modeling and optimization. Society of Petroleum Engineers - Arctic and Extreme Environments Conference and Exhibition, 2. <https://doi.org/10.2118/149926-MS>.
- [85] Belhaj, H., Zaman, M. S., Lay, T (2009). Economical feasibility study of abandoned oilfields utilizing smart modeling approach: case study. OMAE 2009, Offshore Geotechnics - Petroleum Technology, 7, 703-711.
- [86] Lopez, B., Aguilera, R (2015). Physics-based approach for shale gas numerical simulation: Quintuple porosity and gas diffusion from solid kerogen. Proceedings - SPE Annual Technical Conference and Exhibition, 5714-5745. <https://doi.org/10.2118/175115-MS>.
- [87] Bao, X., Chen, Z. J., Wei, Y., Sun, J., Dong, C. C., Deng, H., Song, Y (2010). Geostatistical modeling and numerical simulation of the SAGD process: Case study of an Athabasca reservoir with top water and gas thief zones. Society of Petroleum Engineers - Canadian Unconventional Resources and International Petroleum Conference, 2, 1240-1259. <https://doi.org/10.2118/137435-MS>.
- [88] Sarma, P., Chen, W. H., Durlofsky, L. J., Aziz, K (2008). Production Optimization With Adjoint Models Under Nonlinear Control-State Path Inequality Constraints. SPE Journal, 326-339. <https://doi.org/10.2118/99959-MS>.
- [89] Grebenkin, I., Davies, D. R (2012). A novel optimisation algorithm for inflow control valve management. 74th European Association of Geoscientists and Engineers - Conference and Exhibition, Incorporating SPE EUROPEC: Responsibly Securing Natural Resources, 1749-1762. <https://doi.org/10.2118/154472-MS>.
- [90] Stanley, B (2014). Effect of uncertainty in PVT properties on CO<sub>2</sub> EOR. 38th Nigeria Annual International Conference and Exhibition, Africa's Energy Corridor: Opportunities for Oil and Gas Value Maximization Through Integration and Global Approach, 2, 978-991. <https://doi.org/10.2118/172430-MS>.
- [91] Yoon, S., Alghareeb, Z. M., Williams, J. R (2014). Development of reduced-order oil reservoir models using localized DEIM. Proceedings - SPE Annual Technical Conference and Exhibition, 3, 2206-2216. <https://doi.org/10.2118/170741-MS>.
- [92] Czwienzek, B.F., Perez, J. B., Salve, G. J., Martinez-Ramirez, I., Gerardo, V. M., Aguilar, H. R (2009). Integrated production model with stochastic simulation to define teoteco exploitation plan. SPE Latin American and Caribbean Petroleum Engineering Conference Proceedings, 1, 509-525. <https://doi.org/10.2118/121801-MS>.
- [93] Schebetov, A., Rimoldi, A., Piana, M (2010). Quality check of gas-condensate PVT studies and EOS modelling under input data uncertainty. Society of Petroleum Engineers - SPE Russian Oil and Gas Technical Conference and Exhibition, 1, 125-137. <https://doi.org/10.2118/133258-MS>.
- [94] Nguyen, N. T. B., Dang, C. T. Q., Chen, Z., Nghiem, L. X (2015). Effects of lithofacies and reservoir heterogeneity on improved oil recovery processes. Society of Petroleum Engineers - SPE Canada Heavy Oil Technical Conference, 217-233. <https://doi.org/10.2118/174473-MS>.
- [95] Gao, Y., Yuan, R., Li, R., Will, J., Bai, T., Chang, D (2016). An integrated geomechanics-reservoir modeling workflow for hydraulic fracturing optimisation and EUR prediction for a shale gas play in Sichuan Basin. SPE Asia Pacific Hydraulic Fracturing Conference. <https://doi.org/10.2118/181801-MS>.
- [96] Nguyen, N. T. B., Chen, Z., Nghiem, L. X., Dang, C. T. Q., Yang, C (2014). A new approach for optimization and uncertainty assessment of surfactant-polymer flooding. Society of Petroleum Engineers - 30th Abu Dhabi International Petroleum Exhibition and Conference, ADIPEC 2014: Challenges and Opportunities for the Next 30 Years, 5, 3407-3423. <https://doi.org/10.2118/172003-MS>.
- [97] Pan, Y., Hui, M.-H., Narr, W., King, G., Tankersley, G. T. H., Jenkins, S. D., Flodin, E. A., Bateman, P. W., Laidlaw, C., Vo, H. X (2013). Integration of pressure transient data in modeling Tengiz field, Kazakhstan - A new way to characterize fractured reservoirs. Society of Petroleum Engineers - SPE Western Regional / Pacific Section AAPG Joint Technical Conference: Energy and the Environment Working Together for the Future, 95-107. <https://doi.org/10.2118/165322-PA>.
- [98] Perrone, A., Della Rossa, E (2015). Optimizing reservoir life-cycle production under uncertainty: A robust ensemble-based methodology. Society of Petroleum Engineers - SPE Reservoir Characterisation and Simulation Conference and Exhibition, 967-980. <https://doi.org/10.2118/175570-MS>.
- [99] Revana, K., Erdogan, H. M (2007). Optimization of cyclic steam stimulation under uncertainty. SPE Hydrocarbon Economics and Evaluation Symposium, 203-213. <https://doi.org/10.2118/107949-MS>.
- [100] Hui, M.-H., Mallison, B., Lim, B.K.-T (2008). An innovative workflow to model fractures in a giant carbonate reservoir. International Petroleum Technology Conference, 3, 2047-2061. <https://doi.org/10.3997/2214-4609-pdb.148.iptc12572>.
- [101] Pimenta, R. R. G., Schiozer, D. J., Mello, S. F., Hohendorff Filho, J. C. V (2013). Production strategy selection for a naturally fractured carbonate reservoir from Campos basin. Proceedings of the Annual Offshore Technology Conference, 2, 1127-1136. <https://doi.org/10.4043/24412-MS>.

- [102] Welkenhuysen, K., Meyvis, B., Piessens, K (2017). A profitability study of CO<sub>2</sub>-EOR and subsequent CO<sub>2</sub> storage in the North Sea under low oil market prices. *Energy Procedia*, 114, 7060-7069. <https://doi.org/10.1016/j.egypro.2017.03.1848>.
- [103] Johnson, A. E., Bellion, T., Lim, T., Montini, M., Humphrey, A. I (2013). Managing flow assurance uncertainty through stochastic methods and life of field multiphase simulation. BHR Group - 16th International Conference on Multiphase Production Technology, 61-75. BHR-2013-B1.
- [104] Tilke, P. G., Banerjee, R., Halabe, V. B., Balci, B., Thambynayagam, R. K. M (2010). Optimizing well placement planning in the presence of subsurface uncertainty and operational risk tolerance. ECMOR - 12th European Conference on the Mathematics of Oil Recovery. <https://doi.org/10.3997/2214-4609.20144997>.
- [105] Kelkar, M. G., Pochampally, S., Bahar, A., Sharifi, M (2014). Dynamic vs. Static ranking: Comparison and contrast in application to geo-cellular models. *Proceedings - SPE Annual Technical Conference and Exhibition*, 2, 1341-1357. <https://doi.org/10.2118/170682-MS>.
- [106] Zhang, D., Li, H., Chang, H., Yan, G (2008). Non-intrusive stochastic approaches for efficient quantification of uncertainty associated with reservoir simulations. ECMOR - 11th European Conference on the Mathematics of Oil Recovery. <https://doi.org/10.3997/2214-4609.20146406>.
- [107] Sampaio, M. A., Barreto, C. E. A. G., Ravagnani, A. T. F. S. G., Schiozer, D. J (2011). Comparison between smart and conventional wells optimized under economic uncertainty. *Proceedings of the Annual Offshore Technology Conference*, 1, 533-545. <https://doi.org/10.4043/22426-MS>.
- [108] Van Essen, G. M., Van Den Hof, P. M. J., Jansen, J.-D (2013). A Two-Level Strategy to Realize Life-Cycle Production Optimization in an Operational Setting. *SPE Journal*, 18(6), 1057-1066. <https://doi.org/10.2118/149736-PA>.
- [109] Oliveira, F. S. P. A., Gomes, D. M., Cavalcante, J. R., Leitão, H. C (2015). Optimizing steam injection scheduling using analytical models in a probabilistic approach. *Society of Petroleum Engineers - SPE Canada Heavy Oil Technical Conference*, 1024-1034. <https://doi.org/10.2118/174427-MS>.
- [110] Yue, M., Leung, J. Y., Dehghanpour, H (2016). Numerical investigation of limitations and assumptions of analytical transient flow models in tight oil reservoirs. *Journal of Natural Gas Science and Engineering*, 30, 471-486. <https://doi.org/10.1016/j.jngse.2016.01.042>.
- [111] Temizel, C., Purwar, S., Urrutia, K., Abdullayev, A (2015). Real-time optimization of waterflood performance through coupling key performance indicators in intelligent fields. *SPE Digital Energy Conference and Exhibition*, 121-133. <https://doi.org/10.2118/173402-MS>.
- [112] Conejeros, R., Lenoach, B (2007). Effect of uncertainty on 2-phase flow into a horizontal completion. *Journal of Petroleum Science And Engineering*, 58(1-2), 309-324. <https://doi.org/10.1016/j.petrol.2007.02.006>.
- [113] Orodu, O. D., Tang, Z., Anawe, P. A. L (2011). Sidetrack and recompletion risk evaluation - waterflooded reservoir. *Journal of Petroleum Science and Engineering*, 78, 705-718. <https://doi.org/10.1016/j.petrol.2011.08.015>.
- [114] Niz-Velasquez, E., Bagheri, S. R., Van Dorp, J. J., Verlaan, M. L., Jennings, J (2014). Modelling Development of a Thermal Gas/Oil Gravity-Drainage Process in an Extraheavy-Oil Fractured Reservoir. *Journal of Canadian Petroleum Technology*, 53, 234-246. <https://doi.org/10.2118/169031-PA>.
- [115] Yadali Jamaloei, B., Singh, A. R (2015). Impact of formation dilation-recompaction on development of cyclic steam stimulation (CSS) in an unconventional heavy-oil reservoir: Seal's Cadotte case. *SPE Canada Heavy Oil Technical Conference*, 591-600. <https://doi.org/10.2118/174446-MS>.
- [116] Jenkins, A., Fathi, E., Belyadi, F (2017). Stress field behavior induced by hydraulic fracture in shale reservoirs: A practical view on cluster spacing. *Journal of Natural Gas Science and Engineering*, 48, 186-196. <https://doi.org/10.1016/j.jngse.2016.07.064>.
- [117] Alkhatib, A., Babaei, M., King, P. R (2012). Decision making under uncertainty in EOR-applying the least-squares Monte Carlo method in chemical EOR implementation. *74th European Association of Geoscientists and Engineers Conference and Exhibition Incorporating SPE EUROPEC: Responsibly Securing Natural Resources*, 2544-2563. <https://doi.org/10.2118/154467-MS>.
- [118] Ghosh, B., King, P (2013). Optimisation of smart well completion design in the presence of uncertainty. *Society of Petroleum Engineers - SPE Reservoir Characterisation and Simulation Conference and Exhibition: New Approaches in Characterisation and Modelling of Complex Reservoirs*, 2, 724-740. <https://doi.org/10.2118/166008-MS>.
- [119] Ruijian, L. I., Reynolds, A. C., Oliver, D. S (2009). History matching of three-phase flow production data. *SPE Reprint Series*, 328-340. <https://doi.org/10.2118/87336-PA>.
- [120] Zhao, H., Li, Y., Cui, S., Shang, G., Reynolds, A. C., Guo, Z., Li, H. A (2016). History matching and production optimization of water flooding based on a data-driven interwell numerical simulation model. *Journal of Natural Gas Science and Engineering*, 31, 48-66. <https://doi.org/10.1016/j.jngse.2016.02.043>.
- [121] AlSofi, A. M., Blunt, M. J (2014). Polymer flooding design and optimization under uncertainty. *Journal of Petroleum Science and Engineering*, 124, 46-59. <https://doi.org/10.1016/j.petrol.2014.10.014>.

- [122] Dong, Z., Holditch, S. A., Ayers, W. B., Lee, W. J (2014). Probabilistic estimate of global coalbed methane recoverable resources. SPE USA Unconventional Resources Conference, 511-529. <https://doi.org/10.2118/169006-PA>.
- [123] Hui, M.-H., Kamath, J., Narr, W., Gong, B (2007). Realistic characterization and simulation of naturally fractured reservoirs. International Petroleum Technology Conference, 2, 842-852. <https://doi.org/10.3997/2214-4609-pdb.147.iptc11386>.
- [124] White, J. T., Doherty, J. E., Hughes, J. D (2014). Quantifying the predictive consequences of model error with linear subspace analysis. Water Resources Research, 50, 1152-1173. <https://doi.org/10.1002/2013WR014767>.
- [125] Thomas, P., Bratvold, R. B (2015). A real options approach to the gas blowdown decision. Proceedings - SPE Annual Technical Conference and Exhibition, 1983-2002. <https://doi.org/10.2118/174868-MS>.
- [126] Lodoen, O. P., Tjelmeland, H (2010). Bayesian calibration of hydrocarbon reservoir models using an approximate reservoir simulator in the prior specification. Statistical Modelling, 10(1), 89-111. <https://doi.org/10.1177%2F1471082X0801000106>.
- [127] Bieker, H. P., Slupphaug, O., Johansen, T. A (2007). Real-time production optimization of oil and gas production systems: A technology survey, SPE Prod. Oper. 22(4), 382-391. <https://doi.org/10.2118/99446-PA>.
- [128] O'Hagan, A (2006). Bayesian analysis of computer code outputs: A tutorial, Reliab Eng Syst Saf, 91(10-11), 1290-1300. <https://doi.org/10.1016/j.res.2005.11.025>.
- [129] Blanning, R.W (1975). Construction and implementation of metamodelling, Simulation, 24(6), 177-184. <https://doi.org/10.1177%2F003754977502400606>.
- [130] Kleijnen, J. P. C (2009). Kriging metamodeling in simulation: A review, Eur J Oper Res, 192(3), 707-716. <https://doi.org/10.1016/j.ejor.2007.10.013>.
- [131] Horowitz, B., Guimaraes, L. J., Dantas, V., Afonso, S. M (2010). A concurrent efficient global optimization algorithm applied to polymer injection strategies J Pet Sci Eng., 71, 195-204. <https://doi.org/10.1016/j.petrol.2010.02.002>.
- [132] Davis, P. K., Bigelow, J. H (2002). Motivated Metamodels: Synthesis of cause effect reasoning and statistical metamodeling. Rand: Project Air Force. ISBN-10: 0833033190.
- [133] Razavi, S., Tolson, B. A., Burn, D. H (2012). Review of surrogate modeling in water resources. Water Resour Res 48. <https://doi.org/10.1029/2011WR011527>.
- [134] Murphy P., Messenger, D (2014). Proxy Models: Lessons from other areas. Institute and Faculty of Actuaries. <https://www.actuaries.org.uk/documents/a10-proxy-models-lessons-other-areas>.
- [135] Valtrová, M (2009). Computation Aspects of Kriging in Chosen Engineering Problems [PhD thesis]. Czech University of Life Sciences Prague. Faculty of Environmental Sciences.
- [136] Mohaghegh, S (2000). Virtual-intelligence applications in petroleum engineering: Part 1—artificial neural networks J Pet Technol., 52, 64-73. <https://doi.org/10.2118/58046-JPT>.
- [137] Hasani, M., Emami, F (2008). Evaluation of feed-forward backpropagation and radial basis function neural networks in simultaneous kinetic spectrophotometric determination of nitroaniline isomers. Talanta, 75(1), 116-126. <https://doi.org/10.1016/j.talanta.2007.10.038>.
- [138] Box, G. E. P., Draper, N. R (2007). Response Surfaces, Mixtures, and Ridge Analysis (2nd Edition) John Wiley & Sons, New Jersey. ISBN: 978-0-470-05357-7.
- [139] Nasira, G. M., Kumar, A., Kiruba, S (2008). A Comparative Study of Fuzzy Logic with Artificial Neural Networks Algorithms in Clustering. Journal of Computer Applications, 1(4), 6-8. [http://www.jcaksrce.org/upload/49118197\\_vol2i1p2.pdf](http://www.jcaksrce.org/upload/49118197_vol2i1p2.pdf).
- [140] Silva, L. M., Gonçalves, R. M., Lira, M. M. S., PEREIRA, P. S (2013). Modelagem Fuzzy Aplicada na detecção da vulnerabilidade à erosão costeira. Boletim de Ciências Geodésicas, 19, 746-764. <http://dx.doi.org/10.1590/S1982-217020130004000014>.
- [141] Tesfamariam, S., Sadiq, R (2006). Risk-based environmental decision-making using fuzzy analytic hierarchy process (F-AHP) Stochastic Environmental Research and Risk Assessment, 21(1), 35-50. <https://doi.org/10.1007/s00477-006-0042-9>.
- [142] Kalman, R. E (1960). A new approach to linear filtering and prediction problems Transactions of the ASME. Journal of Basic Engineering 82 (Series D), 35-45. <https://doi.org/10.1115/1.3662552>.
- [143] Evensen, G (1994). Sequential data assimilation with a non-linear quasigeostrophic model using Monte-Carlo methods to forecast error statistics. J Geophys Res., 99, 10143-10162. <https://doi.org/10.1029/94JC00572>.
- [144] Evensen, G (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. Ocean Dynamics, 53 (4), 343-367. <https://doi.org/10.1007/s10236-003-0036-9>.
- [145] Nævdal, G., Johnsen, L. M., Aanonsen, S. I., Vefring, E. H (2005). Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter. Journal SPE, 10(1), 66-74. <https://doi.org/10.2118/84372-PA>.
- [146] Chen, Y., Zhang, D (2006). Data Assimilation for Transient Flow in Geologic Formations via Ensemble Kalman Filter. Advances in Water Resources, 29(8), 1107-1122. <https://doi.org/10.1016/j.advwatres.2005.09.007>.

- [147] Emerick, A. A (2012). History Matching and Uncertainty Characterization Using Ensemble-Based Methods. [PhD thesis], University of Tulsa the Graduate School.
- [148] Okenyi, K., Omeke, J (2012). Well Performance Optimization Using Experimental Design Approach. SPE Annual Technical Conference and Exhibition, Abuja, Nigeria. <https://doi.org/10.2118/162973-MS>.
- [149] Craig P. S., Goldstein M., Seheult A. H., Smith, J. A (1996). Bayes linear strategies for history matching of hydrocarbon reservoirs. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (eds) Bayesian Statistics, Clarendon Press, Oxford, UK, 5, 69–95. [http://www2.stat.duke.edu/~fei/samsi/Oct\\_09/GoldCraigSeheSmit1996.pdf](http://www2.stat.duke.edu/~fei/samsi/Oct_09/GoldCraigSeheSmit1996.pdf).
- [150] Gao, G., Zafari, M., Reynolds, A. C (2006). Quantifying uncertainty for the punq-s3 problem in a Bayesian setting with RML and EnKF. SPE Journal, 11(6), 506-515. <https://doi.org/10.2118/93324-PA>.
- [151] Cumming, J. A., Goldstein, M (2009). Bayes linear uncertainty analysis for oil reservoirs based on multiscale computer experiments. In: O'Hagan, A., West, M. (eds) Handbook of Bayesian Analysis Oxford University Press, Oxford, UK.
- [152] Elsheikh, A. H., Jackson, M. D., Laforce, T. C (2012). Bayesian reservoir history matching considering model and parameter uncertainties. Mathematical Geosciences, 44, 515–543. <https://doi.org/10.1007/s11004-012-9397-2>.
- [153] Al-Mudhafar, W (2015). Integrating Bayesian model averaging for uncertainty reduction in permeability modeling. Offshore Technology Conference. <https://doi.org/10.4043/25646-MS>.
- [154] Yang, C., Nghiem, L., Erdle, J., Moinfar, A., Fedutenko, E., Li, H., Card, C (2015). An efficient and practical workflow for probabilistic forecasting of brownfields constrained by historical data. Society of Petroleum Engineers. <https://doi.org/10.2118/175122-MS>.
- [155] Avansi, G. D (2014). Ajuste de Histórico Integrado à Caracterização de Reservatórios de Petróleo [PhD thesis], Universidade de Campinas.
- [156] Moreno, R., Avansi, G. D., Schiozer, D. J., Vernon, I., Goldstein, M., Caiado, C (2018). Emulation of reservoir production forecast considering variation in petrophysical properties. Journal of Petroleum Science and Engineering, 165, 711-725. <https://doi.org/10.1016/j.petrol.2018.02.056>.
- [157] Melanie, M (1998). An Introduction to Genetic Algorithms. The MIT Press, England. ISBN:9780262133166
- [158] Reynolds, A.C., He, N., Chu, L., Oliver, D. S (1996). Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data. SPE Journal, 1, 413–426. CONF-951002-TRN: 96:000704-0052.
- [159] Ghanem, R., Spanos, P (1991). Stochastic finite elements: a spectral approach. Springer, New York, NY.
- [160] Sarma, P (2006). Efficient closed-loop optimal control of petroleum reservoirs under uncertainty [PhD thesis], Stanford University.
- [161] Anifowose, F., Labadin, J., and Abdulraheem, A (2014). Prediction of petroleum reservoir characterization with a stacked generalization ensemble model of support vector machines. Appl. Soft Comput, 26, 483–496. <http://dx.doi.org/10.1016/j.asoc.2014.10.017>.
- [162] Anifowose, F., Labadin, J., and Abdulraheem, A (2015). Improving the Prediction of Petroleum Reservoir Characterization with a stacked Generalization Ensemble Model of Support Vector Machines. Applied Soft Computing, 26, 483–496. <http://dx.doi.org/10.1016/j.asoc.2014.10.017>.
- [163] Guo, Z., Chen, C., Gao, G., Li, R. C., and Liu, C. (2018). Integration of Support Vector Regression with Distributed Gauss-Newton Optimization Method and its Applications to the Uncertainty Assessment of Unconventional Assets. SPE Reservoir Evaluation & Engineering, 1007-10026, SPE-191373-PA.
- [164] Garcia, J., Levy, A., Tung, A., Yang, R., Kaynig-Fittkau, V (2018). Applying Deep Learning to Petroleum Well Data. Harvard University, Harvard John A Paulson School of Engineering and Applied Sciences, 1-9.
- [165] Li, Y., Sun, R., Horne, R (2019). Deep Learning for Well Data History Analysis. SPE Annual Technical Conference and Exhibition, 16p. <https://doi.org/10.2118/196011-MS>.