Pontifícia Universidade Católica
do Rio de Janeiro

Vinicius Mattoso Reis da Silva

Reservoir characterization based on pressure and temperature transient data, using an ensemble-based method

Dissertação de Mestrado

Dissertation presented to the Programa de Pós-Graduação em Engenharia Mecânica of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia Mecânica.

Advisor : Prof. Márcio da Silveira Carvalho
Co-advisor: Dr. Danmer Maza Quinones

Rio de Janeiro
January 2022

Pontifícia Universidade Católica
do Rio de Janeiro

Vinicius Mattoso Reis da Silva

Reservoir characterization based on pressure and temperature transient data, using an ensemble-based method

Dissertation presented to the Programa de Pós-Graduação em Engenharia Mecânica of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia Mecânica. Approved by the Examination Committee.

Prof. Márcio da Silveira Carvalho
Advisor
Departamento de Engenharia Mecânica – PUC-Rio

Dr. Danmer Maza Quinones
Co-advisor
Laboratório de Microhidrodinâmica e Escoamento em Meios Porosos (LMMP) – PUC-Rio

Dra. Priscila Magalhães Ribeiro
Cenpes – Petrobras

Dr. Alexandre Anoze Emerick
Cenpes – Petrobras

Prof. Abelardo Borges Barreto Junior
Departamento de Engenharia Mecânica – PUC-Rio

Rio de Janeiro, January 26, 2022

Vinicius Mattoso Reis da Silva

Bachelor in Mechanical Engineering by Pontifical Catholic University of Rio de Janeiro (PUC-Rio) - 2019.

Bibliographic data

da Silva, Vinicius Mattoso Reis

Reservoir characterization based on pressure and temperature transient data, using an ensemble-based method / Vinicius Mattoso Reis da Silva; advisor: Márcio da Silveira Carvalho; co-advisor: Danmer Maza Quinones. – 2022.

130 f: il.color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Mecânica, 2022.

Inclui bibliografia

1. Engenharia Mecânica – Teses. 2. Simulação não isotérmica. 3. ES-MDA. 4. Problema inverso. 5. Dados acoplados. I. Carvalho, Márcio da Silveira. II. Maza Quinones, Danmer. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Mecânica. IV. Título.

CDD: 621

To my parents and friends, for all support and encouragement

Acknowledgments

I especially would like to demonstrate my gratitude for my mother, Luzana, and my father, João, for all their effort and dedication. They were fundamental foundations for building who I am today, as well as my brothers Luiza and Luccas. I thank my brothers and family for all the comfort and motivation in every turmoil of everyday life.

To my partner, Karine, I thank you for all your understanding regarding the study weekends that were extremely important, but that end up consuming a lot of our free time.

To my advisor, Prof. Marcio Carvalho and my co-adviser Danmer Maza for the great opportunity to work in his group, for the trust and transmission of knowledge.

To all whom I have had the honor of meeting in the Lmmp group during all these years, for all their advice, help, and friendship.

To Petrobras for financing in part the project.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

da Silva, Vinicius Mattoso Reis; Carvalho, Márcio da Silveira (Advisor); Maza Quinones, Danmer (Co-Advisor). **Reservoir characterization based on pressure and temperature transient data, using an ensemble-based method**. Rio de Janeiro, 2022. 130p. Dissertação de Mestrado – Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

Reservoir characterization is an important tool for production/reservoir management. Well tests are commonly used in reservoir characterization and are the only source of dynamic data during the exploration period. These tests typically measure the pressure, rate and temperature responses at a well during controlled production, injection, or static conditions. Generally, only pressure data is post-processed in reservoir characterization. However, considering only pressure data can lead to misinterpretation associated with the neglected thermal effects, causing errors in reservoir properties estimation and consequently inefficient reservoir management. Besides that, pressure data have several noise sources that may compromise the accuracy of test results. Recent results have shown that temperature data can be used to improve reservoir parameter estimation. In this work, the ensemble smoother with multiple data assimilation method (ES-MDA) was applied in synthetic cases created by an in-house non-isothermal reservoir-well flow simulator that considers the Joule-Thomson heating and cooling, adiabatic fluid expansion/compression, conduction, and convection effects in the thermal energy balance equation. The synthetic measured data was obtained by adding gaussian and harmonics noises to the numerical predictions to simulate equipment and tidal effects, respectively. A sensitivity analysis of the effect of the $\mathbf{C}_D$ matrix used for updating parameters of the ES-MDA method on the parameters estimations was carried out. The results show that adding temperature data to the observed data in the history matching improves the estimates of the reservoir parameters, especially for the skin region and reservoir porosity. For the analyses in which the pressure data had the addition of harmonic noise, the inclusion of temperature data also proved to be of great importance for an accurate characterization of the reservoir.

Keywords

Non-isothermal simulation ES-MDA Inverse problem Coupled data

Resumo

da Silva, Vinicius Mattoso Reis; Carvalho, Márcio da Silveira; Maza Quinones, Danmer. **Caracterização de reservatórios com base em dados transientes de pressão e temperatura, utilizando método baseado em conjunto**. Rio de Janeiro, 2022. 130p. Dissertação de Mestrado – Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

A caracterização de reservatórios é uma ferramenta importante para a gestão da produção do mesmo. Testes de poços são comumente usados para caracterizar reservatórios, pois são as únicas fontes de dados dinâmicos durante a etapa de exploração. Esses testes medem as respostas de pressão e temperatura nos poços, dadas condições controladas de produção, injeção ou estática. Geralmente, apenas os dados de pressão são pós-processados para caracterizar o reservatório. Entretanto considerar apenas os dados de pressão podem levar a interpretações errôneas devido ao negligenciamento dos efeitos térmicos, causando erros na estimativa de propriedades do reservatório e, consequentemente, um gerenciamento ineficiente do mesmo. Além disso, os dados de pressão possuem diversas fontes de ruído que podem comprometer a precisão dos resultados dos testes. Trabalhos recentes mostram que o uso de dados de temperatura podem melhorar a estimativa de parâmetros do reservatório. Neste trabalho, o método *ensemble smoother with multiple data assimilation* (ES-MDA) foi aplicado em casos sintéticos criados por um simulador não-isotérmico de fluxo no reservatório-poço que considera o aquecimento de Joule-Thomson e efeitos de resfriamento, expansão / compressão de fluido adiabático, condução e convecção na equação de balanço de energia. Os dados sintéticos medidos foram obtidos adicionando-se ruídos gaussianos e harmônicos aos sinais calculados para simular ruídos nas medições e efeitos de maré, respectivamente. Foi realizada uma análise de sensibilidade da matriz $\mathbf{C}_D$ do método ES-MDA utilizada na atualização dos parâmetros a serem estimados. Os resultados mostram que o acoplamento dos dados de temperatura aos dados de pressão no ajuste de histórico promoveu uma melhora nas estimativas dos parâmetros do reservatório, principalmente para a região de dano e a porosidade do reservatório. Para as análises, nas quais os dados de pressão tiveram a inclusão de ruído harmônico, a adição de dados de temperatura também se mostrou de grande importância para a caracterização precisa do reservatório.

Palavras-chave

Simulação não isotérmica ES-MDA Problema inverso Dados acoplados

Table of contents

List of figures

List of tables

List of Abreviations

Greek Symbols

α_{factor}	Inflation factor
α_{te}	Earth total Thermal diffusive
δ	Roughness
ε_{JT}	Joule-Thompson coefficient
ϕ	Effective porosity
ρ	Density
θ	Time discretization constant
μ	Viscosity
ν	Kinematic viscosity
φ_t^*	Effective adiabatic-expansion coefficient
λ_e	Thermal conductivity of the earth
λ_{an}	Thermal conductivity of material in annulus
λ_{cem}	Thermal conductivity of cement
λ_t	Thermal conductivity of the saturated porous medium

Acronyms

Capital letters

A	Cross-Section area of the pipe
BHP	Bottom hole pressure
BHT	Bottom hole temperature
C	Compressibility
C_p	Heat capacity
C_T	Lumped parameter, thermal storage constant
D	Inside pipe diameter
$EnKF$	Ensemble Kalman Filter
ES	Ensemble Smoother
$ES-MDA$	Ensemble Smoother with Multiple Data Assimilation
H	Total length of welbore
KF	Kalman Filter
L_R	Relaxation distance

MAP	Maximum a posteriori
PTA	Pressure transient analysis
Q	Flow rate
Re	Reynold's number
T	Temperature, K
T_{ei}	Earth Temperature distribution
T_{ext}	Temperature difference between the T^o and the g_G
TTA	Temperature transient analysis
U	Internal Energy
U_t	Overall heat transfer coefficient

Lowercase

f_D	Dimensionless function
g	Gravity acceleration
g_G	Geothermal Gradient
h	Enthalpy
p	Pressure
r, z	Radial and vertical coordinate
r_{ti}	Inside tubbing radius
r_{to}	Outside tubbing radius
r_{ci}	Inside casing radius
r_{co}	Outside casing radius
r_{wb}	Wellbore radius
s	Saturation, fraction
t	Time, seconds
t_D	Dimensionless time
$\mathbf{v}$	Velocity vector

Subscript

e	Energy source term
JT	Joule-Thompson
i	Initial Conditions
m	Component
o	Oil
p	constant pressure or isobaric
r	Rock
s	Mass source term
sf	Sand face
t	Total
T	Constant Temperature or isotherm
w	Water
wb	Wellbore

"All models are wrong but some are usefull."

George E.P. Box .

1
Introduction

Oil exploration and production still have great relevance for industry and the world economy, as it is the supplier of one of the leading energy sources. Moved by the interest to increase the financial return or even to evaluate the viability of new projects, a good representation of the reservoir is necessary. A realistic representation of an oil field is essential to minimize business risks, since oil production is an activity that requires significant investments.

Reservoir engineering teams play a vital role in reservoir modeling as they carry out well tests. In the exploratory phase of a project, these tests are essential because they provide the dynamic data of bottom-hole pressure (BHP) and bottom-hole temperature (BHT) that are used to estimate reservoir properties.

Until recently, pressure and rate data were used in the oil industry to characterize a reservoir. However, recent works have shown that temperature data can improve reservoir characterization. Sui et al.[13] proposed a multilayer testing method that uses pressure and temperature transient data to investigate the formation properties in commingled multilayer reservoirs. Onur and Palabiyik [14] estimated reservoir properties, such as porosity and Joule-Thomson coefficient, making a history matching using temperature data. Using analytical solutions, Mao and Zeidouni[15] showed that temperature data allows the separate determination of damage zone radius and its permeability. Galvão et al.[[11],[12]] showed that considering the isothermal flow hypothesis can lead to misinterpretation. Panini and Onur [16] also showed that the bottom-hole temperature can be used for reservoir characterization.

Well tests consist of pressure and temperature data acquisition at the bottom of a well during production at a controlled flow. After that, the production is stopped, and data acquisition continues for a certain time. Usually, the tests can take a few hours or days, but there are situations where a test has been performed for several months to obtain more information about the reservoir [17],[18],[19],[20]. When the test is done for a long period, it is called an extended well test (EWT).

Dynamic data obtained from transient tests are used to determine the characteristics of the reservoir. Then, it is compared with a model obtained

from the geology and geophysics teams to confirm or correct the reservoir modeling. An appropriate reservoir modeling helps predict its behavior, thus contributing to field management and assisting the decision making process.

The process of incorporating dynamic data into a numerical model that represents a reservoir is known as history matching (HM). It is an ill-posed inverse problem since infinite combinations can lead to suitable adjustments to the observed data.

History matching can be approached in two different ways: deterministic and probabilistic. The first one updates the analysis parameters until a good enough match between the observed data and the numerical model solution is obtained. This approach can be found in recent works, such as da Silva [21] and Gonçalves[22].

Several models (solutions) are created in the second approach, resulting in several simulated curves. In this case, the variables' uncertainties are quantified, originating a distribution of the probability of occurrences. In this work, the probabilistic approach is used for history matching. Different works using this approach are found in the literature, such as Emerick [23],[24], Silva [4], Silva [25], and Xu and Forouzanfar [26].

In the present work, the flow in the reservoir is considered non-isothermal. Synthetic sets of data based on the problem's solution for a given set of parameters are used as observed data for the history matching. The Ensemble Smoother with Multiple Data Assimilation (ES-MDA), proposed by Emerick and Reynolds in 2013 [23], is used to characterize the reservoir and quantify the uncertainties of the reservoir properties.

1.1 Literature review

1.1.1 Well Test

According to the book "Pressure Transient Formation and Well Testing Convolution, Deconvolution and Nonlinear Estimation"[2], the objectives of well testing are to obtain reservoir properties, identify the reservoir's capacity, detect flow boundaries, if they exist, and determine the well productivity. The extracted data during well testing are usually the pressure and volumetric flow rate. It is one of the most powerful tools for determining well and reservoir parameters under dynamic conditions.

Well tests were introduced around 1920 by the Johnston brothers, and their first use was to obtain the average pressure of the reservoir and also

collect fluid samples[2].

Generally, the test consists of transient pressure data acquisition in the well, which is set to produce at a constant flow rate Q during a production time (tp). This phase of the test is called Drawdown (DD). After this production time, the well is closed and the data acquisition continues for a specific period of time. This second phase is called Build-Up (BU). Figure 1.1 shows the volumetric flow rate behavior during a well test.

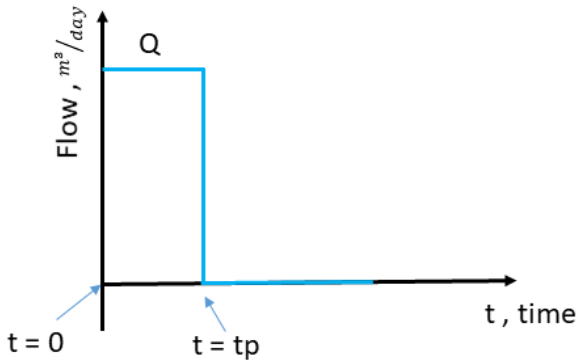

Figure 1.1: Flow rate behavior in a typical well test. During the drawdown period, the flow rate is set to a constant value (Q), and along the buildup period, the flow rate is set to zero.

In 1933, Moore et al.[1] used a history matching approach to estimate the permeability of an infinite-acting radial reservoir. It is noteworthy that this model was adjusted manually by a trial and error procedure. The model did not consider effects such as skin, which was only introduced in 1953 by Van Everdingen et al.[27], neither storage effects introduced in 1949 by van Everdingen and Hurst[28].

Figure 1.2 shows that despite its short test time, when the well is producing at a constant flow rate, there is a pressure drop, and then, during the buildup period, the pressure recovers to values close to the initial pressure.

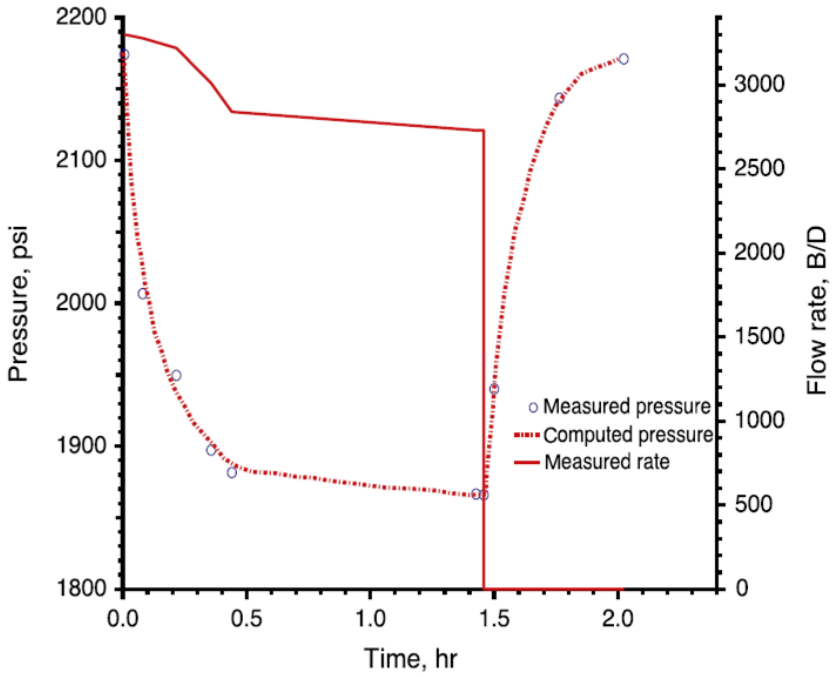

Figure 1.2: Pressure data matching result obtained by Moore et al.[1], extracted from the book " Pressure Transient Formation and Well Testing Convolution, Deconvolution and Nonlinear Estimation"[2].

Reservoir models became more and more complex, as time passed, thus requiring new techniques to analyze the transient pressure data. Gringarten, A. [9] made a summary table of State of the Art evolution about well test analysis from its beginnings in the 1950s until the early 2000s. Table 1.1 was extracted from Gringarten, A.[9].

With the evolution of well-monitoring systems and sensors, temperature information, which was not usually used, began to be studied by several authors, including Duru and Horne[29], Sidorova[30], and Onur et al. [14],[3],[31]. Therefore, driving the creation of new physical models that no longer considered the hypothesis of isothermal flow.

Figure 1.3 is an adaptation of figures presented by Onur and Cinar (2017)[3]. Pressure and temperature data measured at the sandface position are presented for a single-phase flow (oil). Considering the hypotheses adopted by the authors, the behavior of temperature is opposite to that of pressure. While the pressure falls during the production period, the temperature slightly increases. When the buildup period starts, the pressure increases while the temperature decreases.

Date	Interpretation Method	Tools	Emphasis
50s	Straight lines	Laplace transform	Homogeneous reservoir behavior
Late 60s Early 70s	Pressure type-curve analysis	Green's functions	Near-wellbore effects
Late 70s	Type curves with independent variables	Integrated methodology Stehfest algorithm	Dual-porosity behavior
Early 80s	Derivatives	Computerized analysis	Heterogeneous reservoir behavior and boundaries
90s		Computer-aided analysis downhole rate measurements integration with interpretation models from other data	Multilayered reservoir
Early 00s		Deconvolution	Enhanced radius of investigation boundaries

Table 1.1: Summary of the well test historical evolution extracted from Alain C. Gringarten [9].

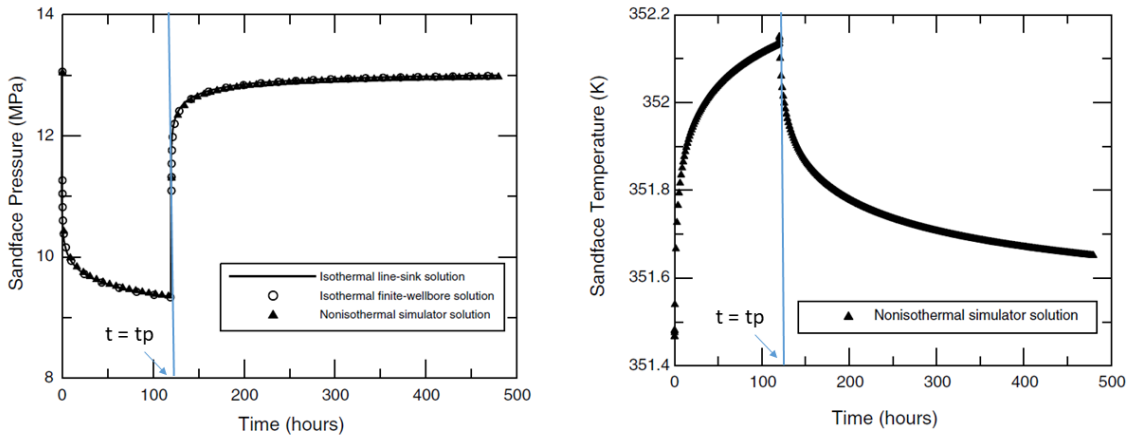

Figure 1.3: Pressure and temperature sandface evolutions adapted from Onur and Cinar [3], with the production time (tp) identified.

Despite the slight temperature variation presented in Fig.1.3, Duru and Horne [32] and Sidorova et al.[30] mention that there are sensors with a resolution of the order of 0.01K and 0.07KPa, which would be enough to capture this slight variation presented. In 2012, Feliciano [10] exhibited information about temperature sensors that have a resolution of 0.001 °C. It confirms the sensors' ability to capture slight variations of temperature accurately.

Single-Point Pressure and Temperature sensor (eletronic)

Maximum Values	15000 psi	150°C
Accuracy	+/- 3psi	+/-0.5°C
Resolution	0.01psi	0.001°C
Lon-term stabiliy	1psi/year	0.1°C/year

Table 1.2: Single-Point Pressure and Temperature sensor (eletronic) informations extracted from Feliciano's work (2012)[10].

1.1.2 Thermal model

According to Valiullin et al.[33], the first measurement of temperature distribution along a well was in 1906, performed by D. Golubyatnikov using a maximal thermometer. This type of thermometer records the maximum temperature to which it has been exposed. However, it was only in the 1930s that temperature measurement gained more attention with the appearance of electric thermometers.

In 1992, Prensky[34] presented an overview of temperature measurements in boreholes. The work mentions that Schlumberger was the first to introduce continuous temperature surveys in 1936[35]. Temperature data along the well made it possible to find the localization of gas inlet and leaking along a well (Dahnov e Dyakjnov[1] ; Hill [37]).

Approximately 20 years after Schlumberger, in 1955, Lawrier [38] presented a mathematical model considering thermal effects in the reservoir. The model considered a 2D reservoir where hot water was injected at a constant temperature and flow rate. In the proposed model, heat transfer occurs only by conduction and convection. Other simplifications were adopted to solve the problem analytically.

Valiullin et al. [36], working on state of the art on temperature logging, mentioned a high-sensitive thermometer with a resolution of hundred parts of the degree appeared in the market (in the 60s). In that same decade, Ramey [39] studied heat transfer near the wellbore, obtaining an analytical approximated solution. With the solutions obtained in this problem, it was possible to estimate the temperature of fluids as a function of depth and time.

In 1965, Chekalyuk[40] developed a single-phase, non-isothermal flow model in porous media, taking into account the Joule-Thomson effect. In 2015, Onur and Palabiyik [14] mentioned that Chekalyuk's original solution did not include the effect of skin factor (S) near-wellbore due to stimulation and damage zone.

[1]This work was not found but was mentioned by Valiulin et al. [36]

Atkinson and Ramey [41], in 1977, introduced a new mathematical model to study heat transfer behavior in fractured and non-fractured porous media. In this work, compressibility and viscous dissipation were not considered to simplify the formulation and obtain analytical solutions.

Garg and Pritchett[42], in the same year, presented a model in which pressure work and viscous dissipation were considered in the energy balance equation for single-phase (liquid water or vapor) or two-phase (liquid water and vapor) flow in porous media. They concluded that considering a steady non-isothermal radial flow, both terms have a small effect on single phase flow (water) and two-phase flow. However, they did not generalize for single-phase vapor flow, because both terms can produce significant variations in this case.

In 1979, Horne and Shinohara[43] analyzed the same problem presented in Ramey's [39] work. They determined the heat-loss rate as a function of fluid flow and properties. They also mentioned that the formulation enabled the evaluation of the wellbore heat loss in many production and injection conditions considering single-phase flow.

In 2007, Ramazanov end Nagimov[44] presented an analytical model that calculates the temperature changes in a saturated porous formation at variable bottom hole pressure. Izgec et al.[45] presented a coupled wellbore/reservoir model. The mass and momentum equations are solved by finite-difference method coupled with a semianalytic heat transfer model, representing the heat transfer in radial and vertical directions.

In 2008, Sui et al.[46] described a model of a multi-layered commingled reservoir coupled with the wellbore. The model is solved numerically in order to provide the transient pressure and temperature behavior. Some simplifications were made to solve the mass balance equation in the wellbore, considering steady-state conditions.

In 2010, App [47] determined the effect of temperature variations caused by the fluid flow on the productivity of wells in conditions where the pressure variations are very high (necessary to reach commercially profitable flows). He concluded that temperature variations could cause an increase in fluid mobility in the region close to the well. He made a numerical study to evaluate the reservoir behavior, considering the Joule-Thomson effect of a 1D, non-isothermal, and high-pressure reservoir. As a result, App showed that the skin effect and the reservoir productivity are underestimated if the oil viscosity is not modeled, considering its variation with pressure.

In the same year, Duru and Horne [29] developed a non-isothermal model in porous medium. The model considers the compressibility, viscous dissipation effects, Joule-Thomson, and adiabatic expansion phenomena. They solved

the reservoir model equations using a semianalytical method called Operator Splitting.

Later, in 2015, Sidorova et al. [30] showed how the temperature transient response depends on the type of testing operations. They mention that two main types of well tests influence temperature measurements: exploration or appraisal well test (new well) and production test (in existing well). The temperature response can show the effect of thermal contamination by mud circulation, fluid losses, and cementing during exploration or appraisal well test. However, during the production well test, the temperature response is affected mainly by geothermal gradient, Joule-Thomson, and adiabatic effects.

In 2016, Onur and Cinar[5] presented an analytical solution for the temperature evolution at the sandface of the reservoir during the flow period. The pressure diffusion equation and the thermal energy balance were decoupled, assuming that the effect of temperature changes in the pressure data can be despised.

In 2017, Onur et al.[31] proposed a semi-analytical solution of the well-reservoir coupled thermal model. They simplify the conservation equations of mass, momentum, and energy that describe the flow in the well, previously presented by Ulker (2016). Here, they neglect the wellbore-temperature gradient $\partial T/\partial z$. The work proposed a methodology to analyze the transient temperature during the drawdown and the buildup periods considering the Joule-Thomson effect, adiabatic expansion/compression, and heat transfer by conduction and convection.

In 2018, Galvão et al.[48] proposed an analytical approach to study the impact on the transient responses of a reservoir well system with the hypothesis of non-isothermal flow. They concluded that when the isothermal flow hypothesis is assumed, errors in the interpretations of the well tests can appear. Mainly it occurs at high flow conditions and high transmissibility reservoirs.

Finally, in 2020, Bircan and Onur[49] presented a semi-analytical solution of the coupled wellbore-reservoir thermal model. Considering constant or variable flow rate and vertical or inclined wells. The models' solutions were used for estimating parameters by nonlinear regression based on a gradient-based method.

Temperature transient analysis - TTA

This chapter presents a brief literature review on transient temperature analysis (TTA) to estimate reservoir parameters. At the beginning of the 2000's, temperature transient analysis began to gain relevance, and the number

of works in this area grew significantly.

In 2009, Valiullin et al.[33] obtained analytical solutions for a vertical well, and showed that the temperature data semi-log graphs could be used to determine flow rates and individual parameters in a multilayer formation. In the following year, Ramanazov et al.[50] used a semi-log plot of temperature data from a field to estimate parameters near the wellbore zone.

Onur et al.[8], in 2016, presented new ways to analyze and interpret temperature transient data acquired during the drawdown and buildup periods from a test. They evaluated the impact of the skin region in the temperature derivative analysis and the temperature semi-log analysis.

Panini and Onur (2018) [16] evaluated how different parameters such as permeability and porosity can impact temperature transient data.

1.1.3 Inverse Problem

Before starting the literature review on inverse problem/history matching, some brief definitions are necessary. Marko Vauhkonen et al.[51], defined a forward or direct problem as: *" to find a unique effect of a given cause using an appropriate physical or mathematical model"*. In other words, if we have a system of equations that describe the phenomena, we can predict the system response by solving analytically or numerically those equations for a given set of parameters.

According to Tarantola [52], *"the inverse problem consists of using the actual result of some measurements to infer the values of the parameters that characterize the system"*. Figure 1.4 better illustrates the difference between these two types of problems.

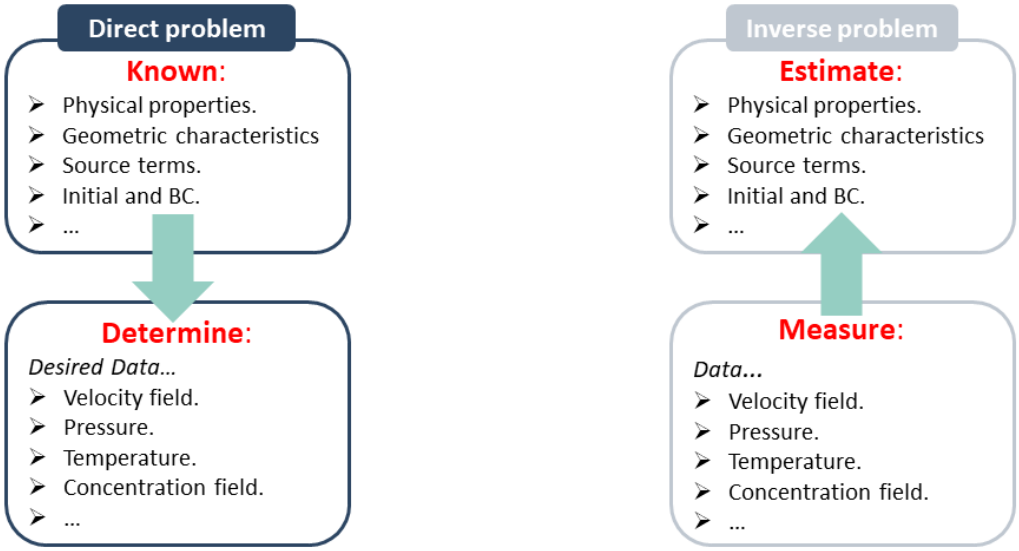

Figure 1.4: Difference between the direct problem and the inverse problems.

Next, a brief review of some works dealing with inverse problems to estimate reservoir parameters are presented. Sui et al. (2008)[13] studied, for synthetic cases, the use of temperature profiles at different well positions and pressure to determine some parameters of the formation. This work *"introduces an entirely new testing approach that uses transient temperature data at multiple locations together with a single-point transient pressure measurement."* They used a Levenberg-Marquardt regression algorithm to estimate the reservoir properties.

In 2010, Duru & Horne [53] solved an inverse problem to obtain an estimate of the permeability and porosity fields in an artificial reservoir, using a combined quasilinear Bayesian inversion method and Ensemble Kalman Filter (EnKF) for data assimilation. As a result, they show that the temperature carries more information about the porosity than the production data (bottom-hole pressure and production rates).

In 2012, Emerick and Reynolds [54] introduced the Ensemble Smoother with Multiple Data Assimilation (ES-MDA). Similar to the Ensemble Smoother (ES), this method used all the available data to make global updates. An iteration of the process of making the history matching and updating the vector of parameters is called assimilation. The ES-MDA method makes interactive assimilation of the observed data, and the number of assimilation times needs to be predefined before the beginning of the method. In the following year, Emerick and Reynolds [24] evaluated the performance of nine different ensemble methods such as EnKF, ES, ES-MDA, and others in terms of data

matching, quantification of uncertainty, and computational cost.

In 2015, Onur e Palabiyik [14] used history matching based on temperature data to estimate reservoir parameters. They compared the Levenberg Marquardt (gradient-based) with the EnKF (non-gradient-based). They show that the Levenberg-Marquardt method leads to estimates closer to the expected values.

In 2016, Silva [4] used the ES-MDA with production data to perform history matching and optimize an oil field production. This work was developed to assist the management of a field during its productive life, considering a closed-loop.

In 2017, Xu and Forouzanfar [26] used wellbore temperature profile, the temperature profile of the reservoir adjacent to the wellbore, and the flowing bottom hole pressure (BHP) of the well at the reference depth, as observed data to use the ES-MDA method. The work was focus on the characterization of a multilayer reservoir assuming a high degree of vertical heterogeniety and homogeneous permeability into the radial direction of each layer. As a result, they show that temperature data significantly influence the characterization of the permeability and porosity of the reservoir layers. Using the temperature data within the ES-MDA significantly reduces the uncertainty in the reservoir parameter estimation.

In 2018, Panini and Onur [16] considered a radial composite reservoir model to estimate reservoir properties based on pressure and temperature data using the Levenberg-Marquardt method. In 2019, Panini et al. [55] also used history matching based on pressure and temperature data. However, in this work, they used the semi-log and log-log analysis to provide the initial guess of the parameters and used the nonlinear regression built-on a gradient-based maximum likelihood estimation (MLE) method to make the estimations.

In 2020, Gonçalves [22] made a parameter estimation using pressure or temperature data, considering three methods: Newton, BFGS associated with a unidirectional search and the variation of the Simplex method called Nelder-Mead Simplex.

1.2 Dissertation goals

The main goal of this work was to use a non-isothermal model to estimate the parameters from a radial heterogeneous reservoir using sandface pressure and/or sandface temperature using an ensemble-based method.

In the literature, there is work using the ES-MDA with temperature data, but the reservoir model adopted has a radial homogenous configuration, and

there are no evaluations about the Cd matrix from the ES-MDA. In order to achieve the main goal of this work, and also fill the literature blanks, this work was divided into two parts. The first part presents the development of a non-isothermal flow simulator to solve the direct problem and calculate the pressure and temperature transient response for a group of reservoir parameters and flow conditions.

In the second part, the ES-MDA method is implemented. Analyses were carried out to evaluate the use of combined pressure and temperature observed data, and the effect of C_D matrix on the parameters estimation accuracy. A comparison was made between cases considering only pressure data and combined pressure and temperature data to evaluate which case made a more accurate parameter estimation and reduced the parameters' uncertainties.

As the last goal, an analysis was made to evaluate the combined data's robustness to estimate reservoir properties. The observed data was created by introducing gaussian and harmonics noises to simulate equipment and tidal effects. Different amplitudes of harmonics noise were considered, and the parameter estimation of these cases was compared with the analysis that considers only the pressure data.

The contribution of this work was the development of computational techniques that improve parameter estimates obtained through production tests, taking into account several effects generally neglected in works available in the literature, and used in industrial procedures. Thermal effects, variation of properties with temperature and pressure, the pressure drop along the well, damaged or stimulated intervals in the reservoir are usually ignored. Important highlights of these contributions are the parametric analysis of the adjustment of the covariance matrix to handle pressure and temperature data simultaneously, and the use of the combined data to improve the radial permeability characterization.

1.3 Dissertation outline

This dissertation is composed of five chapters. Chapter 1 presents an introduction of the subject. It also includes a literature review on well test, non-isothermal reservoir-well model, and methods to estimate reservoir parameters considering well tests data. The objectives and scope of work is presented at the end of the chapter.

Chapter 2 presents the mathematical formulation for the non-isothermal model of single-phase flow in reservoir and well, basic concepts for understanding the ensemble-based method, and the mathematical formulation of

the method used here to solve the inverse problem.

Chapter 3 presents a brief explanation of the finite difference method used to solve the direct problem and a brief explanation of the steps of the ES-MDA.

Chapter 4 presents a numerical simulator validation and the direct problem solution for different reservoir configurations analyzed.

Chapter 5 discusses the effect of the configuration of the C_D matrix of the ES-MDA method and the type of observed data (pressure or combined pressure and temperature) on the parameter estimation. The results of the estimations of the parameters for each case are presented, comparing the analysis made with only pressure data and the one made with combined pressure and temperature. The observed data was generated by adding gaussian and harmonic noises into the simulated data to represent equipment and tidal effects. In this last case, the amplitude of the harmonic noise was more significant than the common literature values to evaluate the robustness of the method.

Finally, chapter 6 presents the conclusions of this work and suggestions for future work.

2
Mathematical Formulation

2.1
Introduction

This chapter is organized as follows. First, the mathematical formulation describing the flow and heat transfer in a coupled wellbore-reservoir system is introduced. Second, some essential concepts for understanding the inverse problem solution method based on the Kalman filter are explained.

2.2
Direct problem

Figure 2.1 shows a schematic of a coupled wellbore-reservoir system. Here, some hypotheses are made for the reservoir and wellbore in order to simplify the model.

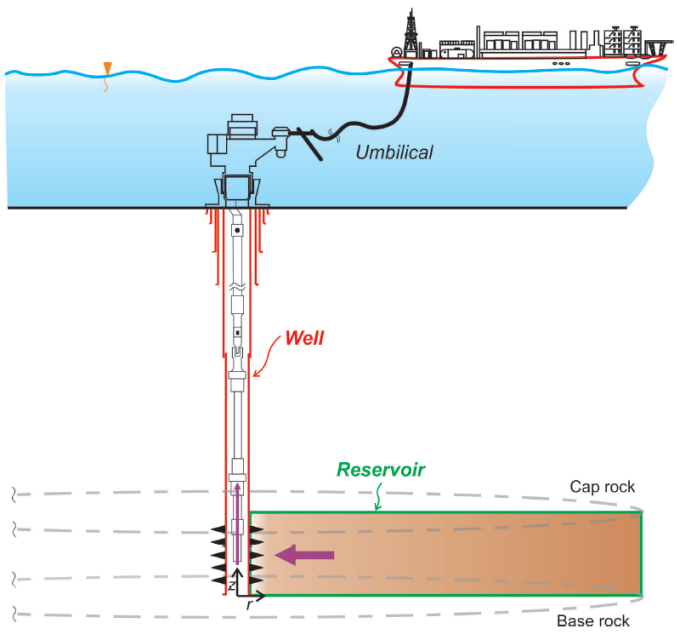

Figure 2.1: Schematic drawing of a radial reservoir cross section.

Some of the hypotheses considered in the present analysis were also assumed in the following works: Onur et al. [5] and [31], Da Silva [21], Gonçalves [22], and De Souza Cardoso[56].

Hypotheses for the reservoir formulation

1. Single-phase oil radial flow with immobile connate water saturation;
2. Reservoir permeability is radially dependent;
3. Oil and water are slightly compressible and immiscible fluids;
4. Fluid flow is governed by Darcy's Law;
5. Reservoir parameters and thermal properties of the fluid (except density and porosity) do not vary with temperature and pressure;
6. Wellbore is vertical and fully penetrates the reservoir.
7. Solid matrix is in local thermal equilibrium with the oil and the connate water, i.e. $T_{solid} = T_{water} = T_{oil} = T$;
8. There is no fluid flow and thermal transfer from the cap and base rocks;
9. Capillary effects are negligible;

Hypotheses for the wellbore formulation

1. Axial flow of slightly compressible single-phase fluid;
2. Heat transfer to the surroundings occurs due to radial diffusion. There is no axial heat diffusion;
3. In the presence of tubing, the casing/tubing annulus is filled with an insulation material;
4. The model takes into account the Joule-Thomson and gravity effects;
5. The pressure and temperature are initially in equilibrium with the geostatic and geothermal gradient;
6. Density is a function of temperature and pressure. Other fluid properties are constant;
7. Wellbore materials have constant thermal conductivities.

2.2.1 Mass conservation equations

Reservoir

To simplify the notation from the mass conservation: the water phase will be represented by the subscript w, and the subscript o will represent the oil phase($m = o, w$). Mass conservation equation can be written as:

$$\phi s_m \frac{\partial}{\partial t}(\rho_m) + s_m \rho_m \frac{\partial}{\partial t}(\phi) + \rho_m \nabla \cdot (\mathbf{v}_m) + \mathbf{v}_m \cdot \nabla(\rho_m) = 0. \tag{2-1}$$

where ϕ, ρ_m, s_m, and $\mathbf{v}_m$ are the porosity, saturation, viscosity, and the velocity of each phase, respectively.

It is assumed that the density and the porosity vary with pressure and temperature. Therefore, the isothermal compressibility and the isobaric thermal expansion coefficients need to be defined for the rock formation and fluids.

$$C_m = \frac{1}{\rho}\left(\frac{\partial \rho_m}{\partial p}\right)\bigg|_T, \tag{2-2a}$$

$$C_r = \frac{1}{\phi}\left(\frac{\partial \phi}{\partial p}\right)\bigg|_T, \tag{2-2b}$$

$$\beta_m = -\frac{1}{\rho}\left(\frac{\partial \rho_m}{\partial T}\right)\bigg|_p, \tag{2-3a}$$

$$\beta_r = -\frac{1}{\phi}\left(\frac{\partial \phi}{\partial T}\right)\bigg|_p. \tag{2-3b}$$

Based on the definition of isothermal compressibility (2-2b) and thermal expansion (2-3b) of the rock, the porosity can be modeled by equation (2-4). Unlike App (2010) [47], the model presented here considers a slightly deformable porous media.

$$\phi(r,t) = \phi^{ini}(r,t)\ exp(C_r \Delta p^r - \beta_r \Delta T^r). \tag{2-4}$$

Equations of state are obtained by using the definitions given in equations (2-2a and 2-3a), in which the density of each fluid phase can be represented as a function of pressure and temperature:

$$\rho_m(r,t) = \rho_m^{ini}(r,t)\ exp(\mathbf{C}_m \triangle p - \beta_m \triangle T). \tag{2-5}$$

Considering equation (2-2a) to equation (2-5), equation (2-1) is rewritten

as:

$$\phi\left(C_m\frac{\partial p}{\partial t}-\beta_m\frac{\partial T}{\partial t}\right)+\nabla\cdot(\mathbf{v}_m)+\frac{1}{\rho_m}\mathbf{v}_m\cdot\nabla(\rho_m)=0. \tag{2-6}$$

Combining the mass balance equation (2-6), for the oil and water phases and assuming irreducible water saturation ($\mathbf{v}_w = 0$), the total mass balance equation becomes:

$$\phi(C_t\frac{\partial p}{\partial t}-\beta_t\frac{\partial T}{\partial t})+\nabla\cdot(\mathbf{v}_o)+\frac{1}{\rho_o}\mathbf{v}_o\cdot\nabla(\rho_o)=0, \tag{2-7}$$

C_t and β_t are defined as:

$$C_t = C_r + s_w C_w + s_o C_o\ , \tag{2-8a}$$

$$\beta_t = \beta_r + s_w\beta_w + s_o\beta_o. \tag{2-8b}$$

Since the model considers the flow to be radial, the oil velocity is written as:

$$v_{ro}=-\frac{K}{\mu_o}\frac{\partial p}{\partial r}. \tag{2-9}$$

Replacing Darcy's velocity in equation (2-7), and considering only radial flow, the final expression of the mass balance is given by:

$$\phi\left(C_t\frac{\partial p}{\partial t}-\beta_t\frac{\partial T}{\partial t}\right)=\frac{1}{r}\left[\frac{\partial}{\partial r}(rv_{ro})+rv_{ro}\left(C_o\frac{\partial p}{\partial r}\right)-rv_{ro}\left(\beta_o\frac{\partial T}{\partial t}\right)\right]. \tag{2-10}$$

Wellbore

The mass balance for the wellbore was proposed by Ulker (2016)[57] and is written as:

$$\frac{\partial p^{wb}}{\partial t}+\frac{Q}{A}\frac{\partial p^{wb}}{\partial z}-\frac{\beta_o}{C_o}\frac{\partial T^{wb}}{\partial t}-\frac{Q\beta_o}{AC_o}\frac{\partial T^{wb}}{\partial z}+\frac{1}{AC_o}\frac{\partial Q}{\partial z}=0. \tag{2-11}$$

The volumetric flow rate $Q(z,t)$ is the velocity of the oil phase inside the well (v_o^{wb}) multiplied by the cross-section area of the well (A).

$$Q(z,t)=Av_o^{wb}. \tag{2-12}$$

The superscript wb means that variables are evaluated in the wellbore, and z

corresponds to the axial coordinate along the well that starts at the bottom hole.

2.2.2 Energy conservation equations

Reservoir

With appropriate use of thermodynamic relations, considerations of local thermal equilibrium between the fluid phase and the rock formation, and including Joule-Thomson effect (ε_{JTo}), the energy conservation equation is written as:

$$\frac{\partial T}{\partial t} + u_{co}(r,t)\frac{\partial T}{\partial r} - \frac{1}{r}\frac{\partial}{\partial r}\left(r\alpha_t \frac{\partial T}{\partial r}\right) - \varphi_t^* \frac{\partial p}{\partial t} - u_{co}(r,t)\varepsilon_{JTo}\frac{\partial p}{\partial r} = 0. \quad (2\text{-}13)$$

Other works use similar formulations, such as App(2010)[47] and Duru and Horne (2010)[53]. All of them are derived from Barenblatt et al. (1989)[58]. In equation (2-13), the terms φ_t^* , $u_{co}(r,t)$, and α_t represent the effective adiabatic-expansion coefficient of the fluid-saturated porous medium, the velocity function of heat transfer, and thermal diffusivity, respectively. They are defined as follow:

$$\varphi_t^* = \frac{(\rho C_p \varphi)_t + \phi p C_r}{(\rho C_p)_t + \phi p \beta_r}, \quad (2\text{-}14)$$

$$u_{co}(r,t) = \frac{\rho_o C_{po}}{(\rho C_p)_t + \phi p \beta_r} v_{ro} = C_{pro} v_{ro}, \quad (2\text{-}15)$$

$$\alpha_t = \frac{\lambda t}{(\rho C_p)_t + \phi p \beta_r}. \quad (2\text{-}16)$$

The adiabatic-expansion coefficient of the reservoir system and the volumetric-heat capacity of the fluid-saturated rock are defined as:

$$(\rho C_p \varphi)_t = \phi(s_o \rho_o C_{po} \varphi_o + s_w \rho_w C_{pw} \varphi_w), \quad (2\text{-}17)$$

$$(\rho C_p)_t = \phi(s_o \rho_o C_{po} \varphi_o + s_w \rho_w C_{pw} \varphi_w) + (1-\phi)\rho_r C_{pr}. \quad (2\text{-}18)$$

Moreover, the adiabatic-thermal expansion of each fluid phase (φ_m) that can be written as a function of Joule-Thomson coefficient, Equation(2-19a, 2-19b), and the effective thermal conductivity of the fluid-saturated rock (λ_t), Equation(2-20), are defined as:

$$\varphi_o = \varepsilon_{JTo} + \frac{1}{(\rho_o C_{po})}, \quad \text{(2-19a)} \qquad \varphi_w = \varepsilon_{JTw} + \frac{1}{(\rho_w C_{pw})}, \quad \text{(2-19b)}$$

$$\lambda_t = \phi(s_o\lambda_o + s_w\lambda_w) + (1-\phi)\lambda_r. \tag{2-20}$$

Wellbore

In 2003, Hasan et al. [59] proposed an energy balance equation that considered the conductive heat loss to the formation, in addition to the convective energy transport along the well. Following in 2017, Onur et al. [31] made some adaptations to the model, leading to:

$$\begin{aligned}\rho_o A c_{po}(1+C_T)\frac{\partial T^{wb}}{\partial t} &= \rho_o Q C_{po} L_R\Big[T_{ext}(z) - T^{wb}(z,t)\Big] \\ &\quad - \rho_o Q c_{po}\left(\frac{\partial T^{wb}}{\partial z} - \varphi(z,t) + \frac{g\sin(\alpha)}{c_{po}}\right).\end{aligned} \tag{2-21}$$

Where C_T is the dimensionless thermal storage coefficient that represents the ratio between the energy of the wellbore and the energy of the fluid. In a similar way to Onur et al. [31], C_T was set to zero in this work. The term $T_{ext}(z)$ is associated with the geothermal gradient, g_G:

$$T_{ext}(z) = T^o - g_G\ z\ sin(\alpha), \tag{2-22}$$

T^o is the reference initial temperature at the bottom of the well, i.e., $z = 0$ and $t = 0$. $\varphi(z,t)$ is a function that takes into account the Joule-Thomson effect and the kinetic-energy contribution, and it is given by:

$$\varphi(z,t) = \varepsilon_{JTo}\frac{\partial p^{wb}}{\partial z} - \frac{Q}{A^2 C_{po}}\frac{\partial Q}{\partial z}. \tag{2-23}$$

L_R is the relaxation-distance parameter that contains the overall heat-transfer coefficient between the fluid and formation and the well components. It was initially proposed by Ramey in 1962 [39], and then it was adapted by Hasan

et al.[59] and Onur et al.[31]. Following Onur et al.[31], it is defined here as.

$$L_R(t) = \frac{2\pi r_{co} U_t \lambda_e}{\rho_o Q C_{po}[r_{co} U_t f_D(t_D) + \lambda_e]}. \tag{2-24}$$

In equation (2-24), λ_e is the thermal conductivity of the formation surrounding the wellbore, $f_D(t_D)$ is the dimensionless heat-transfer function. In Ramey [39], it was considered a line source after Hasan et al.[59] adapted to the cylindrical source presented by Carslaw and Jaeger (1959), considering values in a range of $10^{-6} < t_D < 10^6$:

$$f_D(t_D) = ln[e^{-0.2t_D} + (1.5 - 0.3719e^{-t_D})\sqrt{t_D}], \tag{2-25}$$

where t_D is the dimensionless time represented by the following equation:

$$t_D = \frac{\alpha_{te}}{r_{co}^2} t, \tag{2-26}$$

α_{te} is the thermal-diffusivity coefficient of the formation adjacent to the wellbore.

In equation (2-24), U_t represents the overall heat-transfer from the wellbore to the adjacent formation. It is defined as a function of geometric parameters and the thermal conductivities of materials that compose the well in a transversal section.

Figure 2.2 is a schematic representation of the system with a zoom from the longitudinal section of a production well and its components: packer, tubing, and casing. The region in green represents the oil, which leaves the reservoir and flows through the production column (CoP).

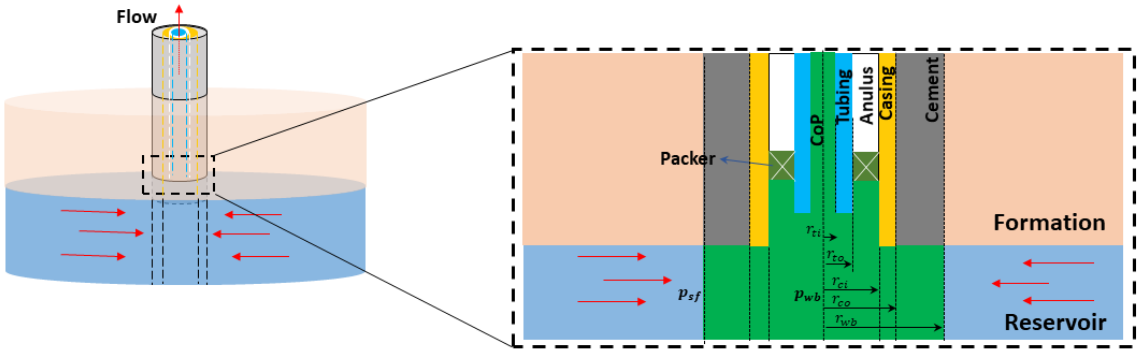

Figure 2.2: Scheme of the wellbore reservoir system, with a zoom in the longitudinal section of the coupled region. Into the zoom, is represented the components of the wellbore, the reservoir, the formation above the reservoir, and in green is the oil that leaves the reservoir and flows through the COP.

The expression of U_t needs to take into account the conduction through the differents walls materials. For example, when the flow occurs only inside the casing, the overall heat-transfer coefficient is given by the following expression:

$$U_t = \frac{1}{r_{ci}}\left[\frac{ln(r_{co}/r_{ci})}{\lambda_{cas}} + \frac{ln(r_{wb}/r_{co})}{\lambda_{cem}}\right]^{-1}. \tag{2-27}$$

According to Hasan and Kabir (2003)[59], the relaxation distance L_R can be understood as an overall resistance for the formation/wellbore system, and the overall heat-transfer coefficient U_t represents the resistance to heat flow in the wellbore.

2.2.3 Momentum Conservation Equation

Wellbore

This work uses the momentum conservation equation for the wellbore presented by Ulker et al. [57], given by:

$$\frac{1}{A}\frac{\partial Q}{\partial t} + \frac{Q}{A^2}\frac{\partial Q}{\partial z} + \frac{1}{\rho_o}\frac{\partial p}{\partial z} + \frac{fQ^2}{2A^2D} + g = 0. \tag{2-28}$$

The equation was developed considering that the flow occurs in the vertical (z) direction and the tube cross-section is rigid.

D is the inside diameter of the pipe, and it can change along the axial (z) direction. f represents the Darcy-Weisbach friction factor, the laminar regime is given by $f = 64/R_e$, in which R_e is the Reynolds number.

$$R_e = \frac{\rho_o Q D}{A \mu_o}. \tag{2-29}$$

In the case of turbulent flow regime ($R_e \geq 2300$), the non-linear Colebrook-White equation is used, Colebrook (1939)[60].

$$\frac{1}{\sqrt{f}} = -2log(\frac{\epsilon}{3.7D} + \frac{2.51}{R_e\sqrt{f}}). \tag{2-30}$$

ϵ is the equivalent well wall roughness.

2.3 Inverse Problem

According to Tarantola [52], the solution of an inverse problem consists of using the response of a system to infer parameters that better describe this system. In this thesis, the inverse problem solved was the estimation of the reservoir parameters that better describe the transient behavior of the bottom hole pressure and temperature during a well test.

Different approaches are used to solve inverse problems, summarized in Fig. 2.3. The approaches can be separated into deterministic and non-deterministic. In the first approach, the goal is to search the reservoir parameters' value that lead to a solution that best fits the observed data. This procedure of adjusting parameters to reproduce the observed data is known in the literature as history matching, and the uncertainty of the evaluated parameters is neglected.

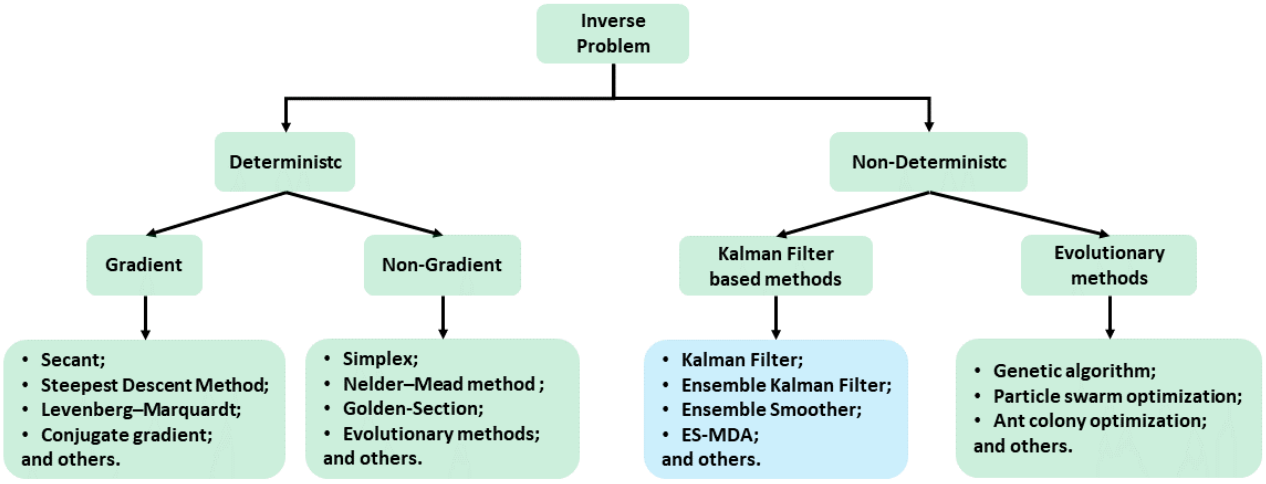

Figure 2.3: Scheme with different methods that can be applied to solve inverse problems.

The second approach seeks to adjust the data and analyze the uncertainties of each predicted parameter. In this group of methods, the methods based on the Kalman filter have been widely used.

In these methods, an ensemble of models is created, where each model consists of a vector with unknown parameters. The history matching is made considering the ensemble and the observed data. As a result of these methods, we have an uncertainty analysis for each unknown parameter and, consequently, an ensemble with the different scenarios of simulated data, considering the final distributions of parameters.

Figure 2.4 shows images illustrating the differences between the two approaches for inverse problem solutions, adapted from Tarantola's book (2005) [52]. Figure (2.4-I) shows the observed data, and it has a specific value for the variable $\mathbf{m}$ where $F(\mathbf{m})$ best fits the observed data. Figure (2.4-II) presents the probability density for F denoted by $\phi(F|\mathbf{m})$. It indicates a range of values for the variable $\mathbf{m}$ that the response produces a good data match with the observed data.

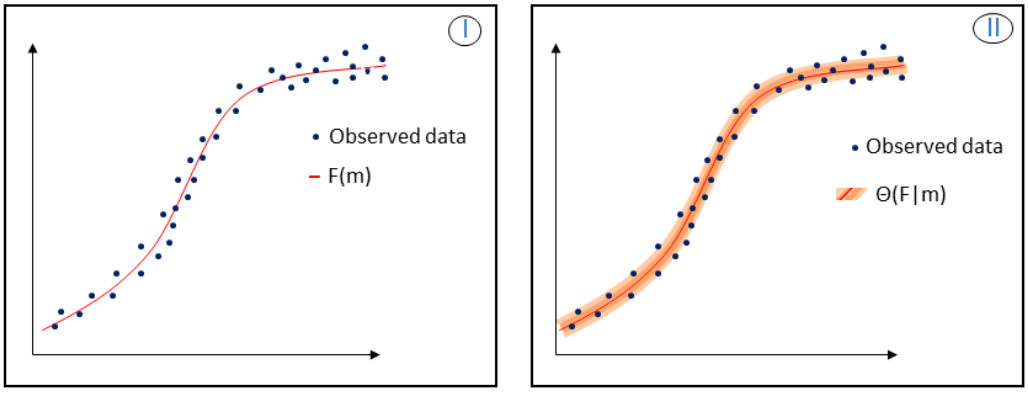

Figure 2.4: I - Data matching in which the uncertainty of the variable **m** can be neglected. II - Data matching in which the uncertainty of the variable **m** is considered.

In this section, some concepts are introduced to better understand the methods based on the Kalman filter.

2.3.1 Basic concepts

Before presenting the mathematical formulation of each method based on Kalman filter, it is necessary to introduce some basic concepts. The first fundamental variable is the vector **m**, also called as model. It consists of a vector with the Nm parameters from the reservoir that we want to estimate. Thus **m** is a column vector of dimension Nm ($\mathbf{m} \in \mathbb{R}^{Nm}$), which represents the number of parameters to be estimated.

$$\mathbf{m} = \begin{bmatrix} m_1 & m_2 & . & . & . & m_j & . & . & . & m_{Nm} \end{bmatrix}^T. \tag{2-31}$$

Considering that $\mathbf{g}(.)$ is a transformation function, we assume that the output of this function, when introduced the argument **m**, is the system's response for the values of parameters defined by **m**. In order words, it is the data predicted by the direct model **d** ($\mathbf{d} \in \mathbb{R}^{Nd}$). The transformation function $\mathbf{g}(\mathbf{m})$ can be considered as the solution of the direct problem for the set of parameters contained in model **m**.

$$\mathbf{d} = \mathbf{g}(\mathbf{m}). \tag{2-32}$$

In our problem, **d** is a vector with the time series of the reservoir response

that can contain the transient response of pressure and/or temperature. This vector can be represented by equation (2-33), where Nd is the number of the discrete outputs from the direct problem.

$$\mathbf{d} = \begin{bmatrix} d_1 & d_2 & . & . & . & d_{Nd} \end{bmatrix}^T. \tag{2-33}$$

Another essential variable is $\mathbf{d}_{obs}$ ($\mathbf{d}_{obs} \in \mathbb{R}^{Nd}$), and it consists of a vector with the observed data (2-34). In other words, it is a vector that contains the observed transient data of the reservoir.

$$\mathbf{d}_{obs} = \begin{bmatrix} d_{obs1} & d_{obs2} & . & . & . & d_{obsNd} \end{bmatrix}^T. \tag{2-34}$$

One way to represent the observed data is by equation (2-35), where an error ϵ is added to the solution of the direct problem using the vector of parameters $\mathbf{m}$. The error ϵ can represent equipment error, oscillations from the tidal effect, or physical phenomena not considered in the model used to describe the reservoir. This error increases the uncertainties associated with data measurements.

$$\mathbf{d}_{obs} = \mathbf{g}(\mathbf{m}) + \epsilon. \tag{2-35}$$

Since the observed data is not a continuous function, a brief review of basic statistical concepts focused on discrete variables is required. The review presented here is based on the book "*Data assimilation: the ensemble Kalman filter*" by Geir Evensen [61] and the publication of Coutinho's thesis[62].

Let $\mathbf{d}$ be a set of random variables ($\mathbf{d}_n$, for $n = 1, ..., N$), the sample mean (μ), is given by:

$$\mu = E[\mathbf{d}] \simeq \overline{\mathbf{d}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{d}_n. \tag{2-36}$$

$E[\mathbf{d}]$ is the expected value, also called as mean value, and represents the best guess as possible of the outcome of $\mathbf{d}$.

The sample variance is another important statistical concept, and it is calculated as follows:

$$\sigma^2 = E[(\mathbf{d} - E[\mathbf{d}])^2] \simeq \overline{(\mathbf{d} - \overline{\mathbf{d}})^2} = \frac{1}{N-1}\sum_{n=1}^{N}(\mathbf{d}_n - \overline{\mathbf{d}})^2. \tag{2-37}$$

Finally, the last statistical concept to be defined is the sample covariance

between two random variables ($\mathbf{d}$ and $\mathbf{m}$). It can be calculated as:

$$Cov(\mathbf{d},\mathbf{m}) = \mathbf{C_{Dm}} = E[(\mathbf{d}-E[\mathbf{d}])(\mathbf{m}-E[\mathbf{m}])] = \frac{1}{N-1}\sum_{n=1}^{N}(\mathbf{d}_n-\overline{\mathbf{d}})(\mathbf{m}_n-\overline{\mathbf{m}}). \tag{2-38}$$

In cases that $\mathbf{d} = \mathbf{m}$ the covariance of the $\mathbf{d}$ and $\mathbf{m}$, is the same as the variance of $\mathbf{d}$.

Another concept many times used in this thesis is an ensemble. It represents the set of vectors of the unknown parameters, and these vectors are also called models.

2.3.2 Bayesian formulation and the objective function of the inverse problem

History matching is performed based on pressure and others measured time series from different reservoir variables. These time series contain a certain level of noise from measurement sensors. The differences between the observed data and model response are related to noise and hypothesis embedded in the model. We call the errors associated with the simplifications as uncertainty of the model.

To achive the objective function, the starting point will be Bayes' theorem. It allows us to write the conditional probability density function (pdf), $f(\mathbf{m}|\mathbf{d}_{obs})$, Equation(2-39a), of the model $\mathbf{m}$ (vector of parameters), given the vector of observed data ($\mathbf{d}_{obs}$). Or even the inverse, the conditional pdf, $f(\mathbf{d}_{obs}|\mathbf{m})$ (Equation(2-39b)), of the vector of observation data, given the vector of model parameters.

$$f(\mathbf{m}|\mathbf{d}_{obs}) = \frac{f(\mathbf{m},\mathbf{d}_{obs})}{f(\mathbf{d}_{obs})}, \tag{2-39a}$$

$$f(\mathbf{d}_{obs}|\mathbf{m}) = \frac{f(\mathbf{m},\mathbf{d}_{obs})}{f(\mathbf{m})}. \tag{2-39b}$$

In the formulas above, $f(\mathbf{m})$ means the prior pdf of the model, $f(\mathbf{d}_{obs})$ is the pdf of the observed data, and $f(\mathbf{m},\mathbf{d}_{obs})$ is the joint pdf, that describes the probability of two events together. Isolating the joint pdf and matching the above equations we obtain:

$$f(\mathbf{m}|\mathbf{d}_{obs}) = \frac{f(\mathbf{d}_{obs}|\mathbf{m})f(\mathbf{m})}{f(\mathbf{d}_{obs})} = \frac{f(\mathbf{d}_{obs}|\mathbf{m})f(\mathbf{m})}{\int_{\mathbf{m}} f(\mathbf{d}_{obs}|\mathbf{m})f(\mathbf{m})d\mathbf{m}} \propto f(\mathbf{d}_{obs}|\mathbf{m})f(\mathbf{m}). \tag{2-40}$$

Equation(2-40) is the Bayes' theorem, which gives the posterior probability of the model parameters considering the observed data, or the conditional probability distribution of the parameters considering a measured data. Another way to refer to $f(\mathbf{d}_{obs}|\mathbf{m})$ is the likelihood function for the model variables, given the data.

$$f(\mathbf{m}|\mathbf{d}_{obs}) \propto L(\mathbf{m}|\mathbf{d}_{obs})f(\mathbf{m}), \tag{2-41}$$

$f(\mathbf{m}|\mathbf{d}_{obs})$ is proportional to the product of $f(\mathbf{m})$ and the likelihood function ($L(\mathbf{m}|\mathbf{d}_{obs})$) for the measurement data. Assuming that the prior parameters pdf and that measurement errors are also a Gaussian distribution, the prior and likelihood function can be defined as:

$$f(\mathbf{m}) \propto \ exp\Bigg(-\frac{1}{2}\ (\mathbf{m}-\mathbf{m}_{pr})^T\ \mathbf{C}_m^{-1}\ (\mathbf{m}-\mathbf{m}_{pr})\Bigg), \tag{2-42}$$

and

$$f(\mathbf{d}_{obs}|\mathbf{m}) \propto \ exp\Bigg(-\frac{1}{2}\ (\mathbf{d}_{obs}-g(\mathbf{m}))^T\ \mathbf{C}_D^{-1}\ (\mathbf{d}_{obs}-g(\mathbf{m}))\Bigg), \tag{2-43}$$

where $\mathbf{m}_{pr}$ is the prior ensemble of the parameter vectors, $\mathbf{C}_m^{-1}$, and $\mathbf{C}_D^{-1}$ are the weighting matrices, for the parameters from model and data, respectively. Those matrices are the inverse of the prior covariance matrix of model parameters $\mathbf{C}_m$ that has a dimension [N_m x N_m], and $\mathbf{C}_D$ is the [N_d x N_d] covariance matrix of measurements errors. Considering that noises from the observed data are independents, and the parameters inside the model ($\mathbf{m}$) are also considered independent, the matrices can be considered diagonal and defined as:

$$\mathbf{C}_D = \begin{bmatrix} \sigma^2_{d_{obs1}} & 0 & \dots & 0 \\ 0 & \sigma^2_{d_{obs2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2_{d_{obsNd}} \end{bmatrix}, \tag{2-44}$$

For the particular case that the model parameters are radial permeability kr and porosity ϕ, $\mathbf{C}_m$ is given by

$$\mathbf{C}_m = \begin{bmatrix} \mathbf{C}_{\phi,\phi} & \mathbf{C}_{kr,\phi} \\ \mathbf{C}_{\phi,kr} & \mathbf{C}_{kr,kr} \end{bmatrix}. \tag{2-45}$$

$\mathbf{C}_{\phi,\phi}$,$\mathbf{C}_{kr,kr}$,and $\mathbf{C}_{\phi,kr}$ are the covariance of the variables from the model. In this work the radial permeability(kr) and the porosity(ϕ) are set independently, so $\mathbf{C}_{\phi,kr} = 0$. Multiplying Equation(2-42) by Equation(2-43).

$$\begin{aligned} f(\mathbf{m}|\mathbf{d}_{obs}) = \propto \ exp\bigg[-\frac{1}{2} \ (\mathbf{d}_{obs} - g(\mathbf{m}))^T \ \mathbf{C}_D^{-1} \ (\mathbf{d}_{obs} - g(\mathbf{m}))\bigg] \\ * \propto \ exp\bigg[-\frac{1}{2} \ (\mathbf{m} - \mathbf{m}_{pr})^T \ \mathbf{C}_m^{-1} \ (\mathbf{m} - \mathbf{m}_{pr})\bigg], \end{aligned} \tag{2-46}$$

reorganizing,

$$\begin{aligned} f(\mathbf{m}|\mathbf{d}_{obs}) = \propto \ exp\bigg(-\frac{1}{2} \ \bigg[(\mathbf{d}_{obs} - g(\mathbf{m}))^T \ \mathbf{C}_D^{-1} \ (\mathbf{d}_{obs} - g(\mathbf{m})) \\ + (\mathbf{m} - \mathbf{m}_{pr})^T \ \mathbf{C}_m^{-1} \ (\mathbf{m} - \mathbf{m}_{pr})\bigg]\bigg). \end{aligned} \tag{2-47}$$

Based on the Equation 2-47, we can say that the object function ($O(\mathbf{m})$) takes account the uncertainty of the model and make history matching with the observed data and can be defined as:

$$O(\mathbf{m}) = O_{model}(\mathbf{m}) + O_{data}(\mathbf{m}). \tag{2-48}$$

The argument of the exponential function in Equation(2-47) is the objective function, in which it is possible to separate the part of adjustment data ($O_{data}(\mathbf{m})$) and model ($O_{model}(\mathbf{m})$).

$$O_{data}(\mathbf{m}) = [(\mathbf{d}_{obs} - g(\mathbf{m}))^T \ \mathbf{C}_D^{-1} \ (\mathbf{d}_{obs} - g(\mathbf{m}))], \tag{2-49}$$

$$O_{model}(\mathbf{m}) = [(\mathbf{m} - \mathbf{m}_{pr})^T \ \mathbf{C}_m^{-1} \ (\mathbf{m} - \mathbf{m}_{pr})]. \tag{2-50}$$

Rewriting Equation(2-47) using the the objective function (eq:2-48)

$$f(\mathbf{m}|\mathbf{d}_{obs}) = \propto \ exp(-O(\mathbf{m})). \tag{2-51}$$

Therefore, to maximize the posteriori distribution of $\mathbf{m}$ given $\mathbf{d}_{obs}$ ($f(\mathbf{m}|\mathbf{d}_{obs})$) is equivalent to minimize the objective function ($O(\mathbf{m})$).

2.3.3 The maximum a posteriori estimate

The maximum a posteriori estimate - MAP ($\mathbf{m}_{MAP}$) is the vector of parameters that minimize the objective function given in Equation(2-48), which is equivalent to maximize the posterior distribution of the model $\mathbf{m}$ given $\mathbf{d}_{obs}$, Equation(2-39a).

$$\mathbf{m}_{MAP} = min\ [O(\mathbf{m})]. \tag{2-52}$$

As mentioned before, the function $\mathbf{g}(\mathbf{m})$ is the predictions obtained by the solution of the direct problem, i.e., the reservoir simulator, which is highly non-linearity. $f(\mathbf{m}|\mathbf{d}_{obs})$ will not be a Gaussian, even if the $f(\mathbf{m})$ is Gaussian. It means that the objective function (Equation(2-48)) will have many local minimums. Thus the MAP estimate will not be unique because the MAP applies strictly only to linear models, with $f(\mathbf{m})$ and measurement errors being Gaussians.

Assuming a linear theoretical relationship between the predicted data and the input parameters, in order to the posteriori $f(\mathbf{m}|\mathbf{d}_{obs})$ be a Gaussian, we can write this relationship as follows:

$$\mathbf{g}(\mathbf{m}) = \mathbf{G}\mathbf{m}. \tag{2-53}$$

Where $\mathbf{G}$ is a matrix [N_d x N_m] that, considering the linear assumption, describes the theoretical relationship between the variables and data. To find the MAP, following the steps in the book "*Inverse theory for petroleum reservoir characterization and history matching*" by Oliver et al.[63], it is necessary to calculate the gradient of the objective function ($O(\mathbf{m})$).

$$\nabla O(\mathbf{m}) = \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{pr}) + G^T\mathbf{C}_D^{-1}(\mathbf{G}\mathbf{m} - \mathbf{d}_{obs}). \tag{2-54}$$

The objective function has the following Hessian matrix,

$$H[O(\mathbf{m})] = \nabla[\nabla O(\mathbf{m})^T] = \mathbf{C}_m^{-1} + \mathbf{G}^T\mathbf{C}_D^{-1}\mathbf{G}. \tag{2-55}$$

The Hessian H can be shown to be a positive definite and hence $\nabla O(\mathbf{m})$ has a unique global minimum which can be found by setting $\nabla O(\mathbf{m}) = 0$ in Equation(2-54).

$$\mathbf{m}_{MAP} = \mathbf{m}_{pr} + \mathbf{C}_m\mathbf{G}^T(\mathbf{C}_D + \mathbf{G}\mathbf{C}_m\mathbf{G}^T)^{-1}(\mathbf{d}_{obs} - \mathbf{G}\mathbf{m}_{pr}). \tag{2-56}$$

The next steps are the expansions and approximations of the terms contained in the MAP equation. The first multiplicative term can be expressed as:

$$\begin{aligned}
\mathbf{C}_m\mathbf{G}^T &= \left(\frac{1}{N_e-1}\sum_{i=1}^{N_e}(\mathbf{m}_i - \overline{\mathbf{m}})(\mathbf{m}_i - \overline{\mathbf{m}})^T\right)\mathbf{G}^T, \\
&= \frac{1}{N_e-1}\sum_{i=1}^{N_e}(\mathbf{m}_i - \overline{\mathbf{m}})(\mathbf{m}_i - \overline{\mathbf{m}})^T\mathbf{G}^T, \\
&= \frac{1}{N_e-1}\sum_{i=1}^{N_e}(\mathbf{m}_i - \overline{\mathbf{m}})(\mathbf{G}m_i - \mathbf{G}\overline{\mathbf{m}})^T, \\
&= \frac{1}{N_e-1}\sum_{i=1}^{N_e}(\mathbf{m}_i - \overline{\mathbf{m}})(\mathbf{d}_i - \overline{\mathbf{d}})^T \approx \mathbf{C}_{mD}.
\end{aligned} \tag{2-57}$$

The multiplicative therm inside the inversion can be rewritten as:

$$\begin{aligned}
\mathbf{G}\mathbf{C}_m\mathbf{G}^T &= \mathbf{G}\left(\frac{1}{N_e-1}\sum_{i=1}^{N_e}(\mathbf{m}_i - \overline{\mathbf{m}})(\mathbf{m}_i - \overline{\mathbf{m}})^T\right)\mathbf{G}^T, \\
&= \frac{1}{N_e-1}\sum_{i=1}^{N_e}\mathbf{G}(\mathbf{m}_i - \overline{\mathbf{m}})(\mathbf{d}_i - \overline{\mathbf{d}})^T, \\
&= \frac{1}{N_e-1}\sum_{i=1}^{N_e}(\mathbf{d}_i - \overline{\mathbf{d}})(\mathbf{d}_i - \overline{\mathbf{d}})^T \approx \mathbf{C}_{DD}.
\end{aligned} \tag{2-58}$$

Therefore, MAP Equation(2-56) can be approximated as:

$$\begin{aligned}
\mathbf{m}^a &= \mathbf{m}_{pr} + \mathbf{C}_{mD}(\mathbf{C}_D + \mathbf{C}_{DD})^{-1}(\mathbf{d}_{obs} - \mathbf{G}\mathbf{m}_{pr}), \\
&= \mathbf{m}_{pr} + \mathbf{C}_{mD}(\mathbf{C}_D + \mathbf{C}_{DD})^{-1}(\mathbf{d}_{obs} - \mathbf{d}).
\end{aligned} \tag{2-59}$$

Due to the approximations made, equation (2-59) can no longer be called MAP. To calculate the inverse matrix was used the function inv(matrix), from the matlab. The following Kalman filter-based methods are optimization methods used to obtain $\mathbf{m}^a$ in equation (2-59).

2.3.4 Kalman Filter-Based Methods

The following methods are based on the Kalman Filter (KF) developed by Kalman in 1960[64]. The KF is a recursive filter for linear systems. In order to extend to a non-linear and high-dimensional problem, Geir Evensen in 1994

introduced the Ensemble Kalman Filter (EnKF)[65]. Commonly, the vector to be updated every new data during the data assimilation in the EnKF contain the parameters to be estimated and the state vector, the state vector is the simulated data [66].

In some cases, the high quantity number of parameters and several different times in the data assimilation process can increase the computational cost of the EnKF. Trying to avoid run the solver of the direct problem many time, Van Leeuwen and Evensen in 1996 [67] introduced the ensemble smoother (ES), which assimilates all the available data simultaneously, producing a global update.

In 2011, Skjervheim et al.[68] compared the EnKF and ES using synthetic and real field data, and concluded that both methods present similar performance and good history matching, which makes it impossible to state which method is better by analyzing the data matching aspects. However, the ES consumed only 10% of the simulation time used by EnKF.

The Ensemble Smoother with Multiple Data Assimilation (ES-MDA), was introduced by Emerick and Reynolds in 2012[23] and, it can be considered as an iterative form of ES. While the ES makes only a single global update, it can produce a non-satisfying result. In order to improve the data matching the ES-MDA method uses a procedure based on multiple global assimilations of the same data with an inflated covariance of the measurement errors. Emerick and Reynolds [24] compared the performance of many ensemble-based methods regarding the quality of the data matches, quantification of uncertainty, and computational cost. The results indicate that the ES-MDA presents better results compared with the EnKF and ES.

Figure 2.5 presentes an adaptation of the scheme, produced by Silva, V. L. S. in 2016 [4], of each ensemble method followed by the mathematical explanation of each one.

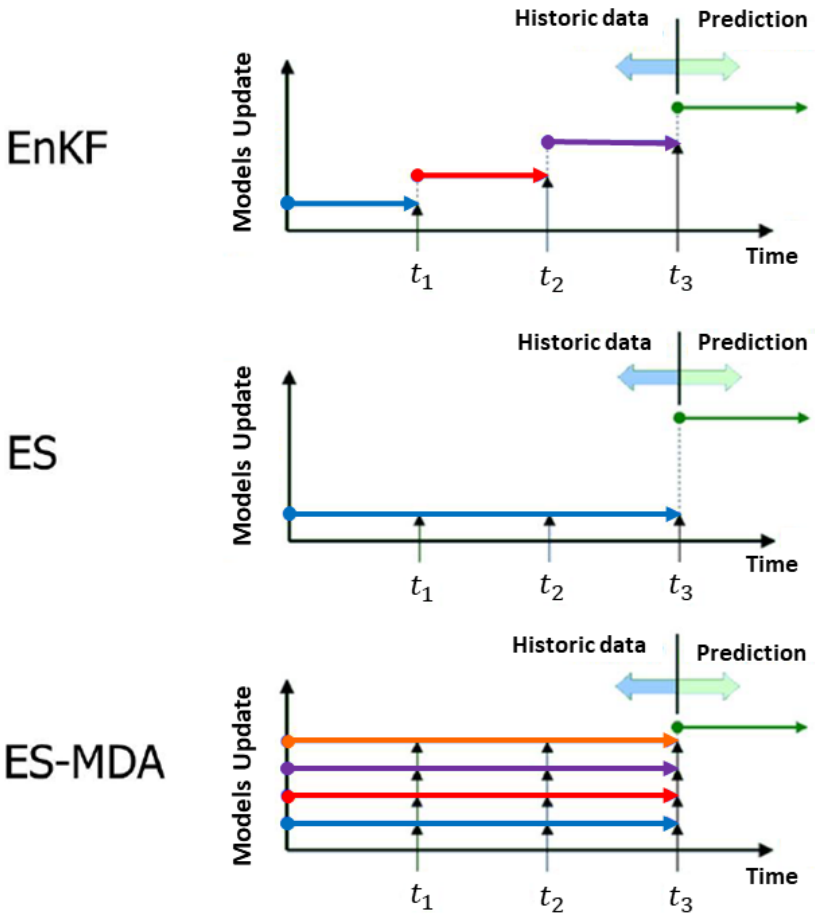

Figure 2.5: This figure is an adaptation of Figure 3.1 from Silva, Vinicius [4], presenting a schematic representation of the EnKF, ES, and ES-MDA methods. Each color represents an update of the model (vector of parameters), the horizontal axis represents the time series of the data. The labels t1, t2, t3 represent different observed data available at that time.

Ensemble Kalman filter - EnKF

In the EnKF, as previously mentioned, the parameters and states are updated. For this, a new vector $\mathbf{y}_j^n$ is created containing the model parameters $\mathbf{m}^n$ and the state $\mathbf{p}_j^n$. Where the superscript n is referent to the ensemble and the subscribed j is referent to the simulation step time.

$$\mathbf{y}_j^n = \begin{bmatrix} \mathbf{m}^n \\ \mathbf{p}_j^n \end{bmatrix}. \tag{2-60}$$

Rewriting Equation(2-59), considering the update of the state and the model instead of considering only the model vector.

$$\mathbf{y}_j^{n+1} = \mathbf{y}_j^n + \tilde{\mathbf{K}}_n(\mathbf{d}_{dobs} - \mathbf{d}_j^n), \ for\ j = 1, ...N_e. \tag{2-61}$$

The term $\tilde{K_n}$ in Equation(2-61) is the Kalman gain matrix. This weighting matrix defines how the difference between the observed data $\mathbf{d}_{obs}$ and the calculated data $\mathbf{d}$, will interfere in the new vector $\mathbf{y}_j^{n+1}$. N_e is the size of the ensemble. Considering the vector $\mathbf{y}$ and the Equation(2-59), the Kalman gain matrix is defined as:

$$\tilde{\mathbf{K}}_n = \mathbf{C}_{YD}(\mathbf{C}_{DD} + \mathbf{C}_D)^{-1}. \tag{2-62}$$

The $\mathbf{C}_{YD}$ is the covariance matrix between the state and model vector $\mathbf{y}$ and the calculated data $\mathbf{d}$, and it can be calculated as:

$$\mathbf{C}_{YD} = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}(\mathbf{y}_i - \overline{\mathbf{y}})(\mathbf{d}_i - \overline{\mathbf{d}})^T. \tag{2-63}$$

Ensemble Smoother - ES

In the ES, the data is not assimilated sequentially in time. Therefore all the available data is used in single global assimilation. The update formula is equivalent to the Equation(2-59).

$$\mathbf{m}_j^{final} = \mathbf{m}_j^{prior} + \mathbf{C}_{MD}(\mathbf{C}_{DD} + \mathbf{C}_D)^{-1}(\mathbf{d}_{obs} - \mathbf{d}_j). \tag{2-64}$$

The terms inside the Equation(2-64) are represented again in order to facilitate their comprehension.

$$\mathbf{C}_{MD} = \frac{1}{N_e - 1}\sum_{j=1}^{N_e}(\mathbf{m}_j - \overline{\mathbf{m}})(\mathbf{d}_j - \overline{\mathbf{d}})^T,$$

$$\mathbf{C}_{DD} = \frac{1}{N_e - 1}\sum_{j=1}^{N_e}(\mathbf{d}_j - \overline{\mathbf{d}})(\mathbf{d}_j - \overline{\mathbf{d}})^T,$$

$$\overline{\mathbf{d}} = \frac{1}{N_e}\sum_{j=1}^{N_e} d_j, \qquad \overline{\mathbf{m}} = \frac{1}{N_e}\sum_{j=1}^{N_e}\mathbf{m}_j.$$

Ensemble Smoother with Multiple Data Assimilation - ES-MDA The ES-MDA method also makes assimilation considering all the available data, similar to the ES method. However, ES-MDA performs the data assimilation

process iteratively, with the number of iterations (N_a) previously defined. Compared with the ES update formula (Equation(2-64)), in ES-MDA there is an addition of the inflation factor (α_{factor}) associated with the $\mathbf{C}_D$ matrix.

$$\mathbf{m}_j^{l+1} = \mathbf{m}_j^l + \mathbf{C}_{MD}^l(\mathbf{C}_{DD}^l + \alpha_{factor}\mathbf{C}_D)^{-1}(\mathbf{d}_{duc}^l - \mathbf{d}_j^l). \tag{2-66}$$

The inflation factor (α_{factor}) must obey the following formation rule:

$$\sum_{l=1}^{N_a} \frac{1}{\alpha_{factor}} = 1. \tag{2-67}$$

There are many combinations of values that obey this formation rule, and it depends on the number of assimilations (N_a) that will be performed. Emerick in 2016 [69] showed that choosing the inflation factor constant and equal to N_a gives a good data matching, especially when considering $N_a = 4$.

Another term not considered in Equation(2-59) is the $\mathbf{d}_{uc}$, a vector of perturbed observed data. Those perturbations are made at each assimilation by the following equation. This procedure tends to reduce the matching problem of a possible outlier that could be generated when the observed data are perturbed in each assimilation.

$$\mathbf{d}_{uc} = \mathbf{d}_{obs} + \sqrt{\alpha_{factor}}\mathbf{C}_D^{1/2} z_d, \tag{2-68}$$

z_d in the equation above is a normal ditribution $N(0, I_{N_d})$, where N_d is the size of the observed data.

3
Numerical Solution

The non-linear differential equations representing the coupled well-reservoir system are discretized in space using a finite difference method on a non-uniform mesh and integrated in time by a second-order implicit scheme. A fixed point iteration method was used to linearize the resulting algebraic equations. The resulting system of algebraic equations is solved by a direct LU method.

3.1
Discretization of the coupled wellbore-reservoir system equation

3.1.1
Discretization of the reservoir equations

The conservations equations [2-10 and 2-13] are non-linear, presenting second-order terms in the reservoir domain. Figure 3.1 illustrates the discretization procedure used for the reservoir domain.

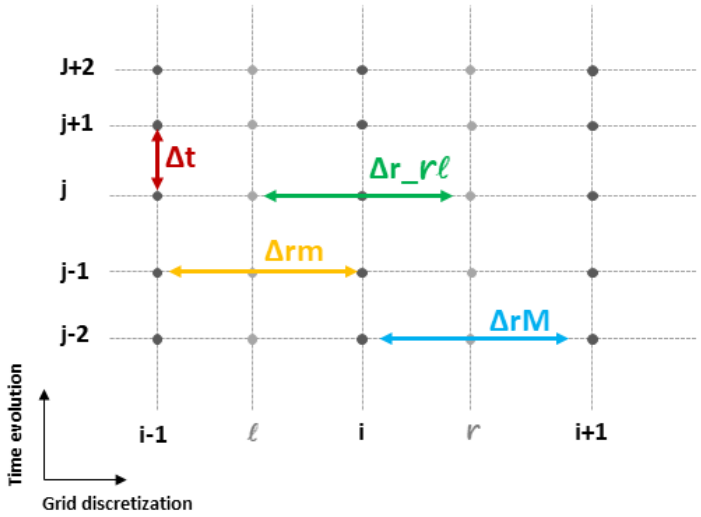

Figure 3.1: Finite Difference scheme for the reservoir discretization.

Following, we present the discretization of mass conservation equation for the reservoir domain. In equation (2-13), to model a heterogeneous medium, the diffusive term is discretized as:

$$\frac{1}{r}\left[\frac{\partial}{\partial r}(rv_{ro})\right] = \frac{D^*}{r_i}\left[\frac{r_r K_r \frac{\partial p}{\partial r}|_r - r_l K_l \frac{\partial p}{\partial r}|_l}{\Delta r_{rl}}\right]. \tag{3-1}$$

Where $D^* = \frac{1}{\phi c_t \mu_o}$, and the subindex r and l are related to the intermediate plan in the right and left of the node, respectively. The harmonic mean was used to calculate the intermediates planes permeability to ensure mass conservation in a non-homogeneous domain.

$$K_l = \frac{2K_{i-1}K_i}{K_i + K_{i-1}}, \tag{3-2a}$$

$$K_r = \frac{2K_{i+1}K_i}{K_{i+1} + K_i}. \tag{3-2b}$$

The discretization of the pressure derivatives terms in the r and l positions follows a central difference approximation.

$$\frac{\partial p}{\partial r}|_l = \frac{p_i - p_{i-1}}{\Delta rm}, \tag{3-3a}$$

$$\frac{\partial p}{\partial r}|_r = \frac{p_{i+1} - p_i}{\Delta rM}. \tag{3-3b}$$

The mass conservation equation of the reservoir (eq:2-10) is discretized as follow:

$$\frac{\partial p}{\partial t} - \frac{\beta_t}{c_t}\frac{\partial T}{\partial t} = \frac{D^*}{r_i}\left[\frac{r_r K_r \frac{\partial p}{\partial r}|_r - r_l K_l \frac{\partial p}{\partial r}|_l}{\Delta r_{rl}}\right] + D^* c_o K_i \left(\left.\frac{\partial p}{\partial r}\right|_j\right)\left(\frac{\partial p}{\partial r}\right) + D^* \beta_o K_i \left(\left.\frac{\partial p}{\partial r}\right|_j\right)\left(\frac{\partial T}{\partial r}\right). \tag{3-4}$$

Expanding the left hand side we have:

$$\frac{\partial p}{\partial t} - \frac{\beta_t}{c_t}\frac{\partial T}{\partial t} = \frac{p_i^{j+1} - p_i^j}{\Delta t} - \left(\frac{\beta_t}{c_t}\right)\frac{T_i^{j+1} - T_i^j}{\Delta t}. \tag{3-5}$$

First term of the right hand side:

$$\frac{D^*}{r_i}\left[\frac{r_r K_r \frac{\partial p}{\partial r}|_r - r_l K_l \frac{\partial p}{\partial r}|_l}{\Delta r_{rl}}\right] = \frac{D^*}{r_i}\left[\theta\left(\frac{r_r K_r \frac{p_{i+1}^{j+1} - p_i^{j+1}}{\Delta rM} - r_l K_l \frac{p_i^{j+1} - p_{i-1}^{j+1}}{\Delta rm}}{\Delta r_{rl}}\right)\right.$$
$$\left. + (1-\theta)\left(\frac{r_r K_r \frac{p_{i+1}^{j} - p_i^{j}}{\Delta rM} - r_l K_l \frac{p_i^{j} - p_{i-1}^{j}}{\Delta rm}}{\Delta r_{rl}}\right)\right]. \tag{3-6}$$

The second term of the right hand side:

$$D^* c_o K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \left(\frac{\partial p}{\partial r} \right) = D^* c_o K_i \frac{p_{i+1}^j - p_{i-1}^j}{\Delta rm + \Delta rM} \left[\theta \frac{p_{i+1}^{j+1} - p_{i-1}^{j+1}}{\Delta rm + \Delta rM} + (1-\theta) \frac{p_{i+1}^j - p_{i-1}^j}{\Delta rm + \Delta rM} \right]. \tag{3-7}$$

The third term of the right hand side:

$$D^* \beta_o K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \left(\frac{\partial T}{\partial r} \right) = D^* \beta_o K_i \frac{p_{i+1}^j - p_{i-1}^j}{\Delta rm + \Delta rM} \left[\theta \frac{T_{i+1}^{j+1} - T_{i-1}^{j+1}}{\Delta rm + \Delta rM} + (1-\theta) \frac{T_{i+1}^j - T_{i-1}^j}{\Delta rm + \Delta rM} \right]. \tag{3-8}$$

The energy conservation equation (2-13) of the reservoir was linearized, evaluating the Darcy velocity in the previous time step:

$$\frac{\partial T}{\partial t} - \varphi_t^* \frac{\partial p}{\partial t} = C_{pRo} K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \frac{\partial T}{\partial r} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \alpha_t \frac{\partial T}{\partial r} \right) - C_{pRo} K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \varepsilon_{JTo} \frac{\partial p}{\partial r}. \tag{3-9}$$

In appendix A, equation (3-9) is fully discretized. In the next section, the discretization of the equations that describe the well is presented.

3.1.2 Discretization of the wellbore equations

This section presents the discretization of the mass conservation equation for the well as an example of the discretization procedure used in this work. Figure 3.2 contains the diagram of the discretization of the nodes in a segment of the well. Each point n in the well domain represents a cylindrical well segment.

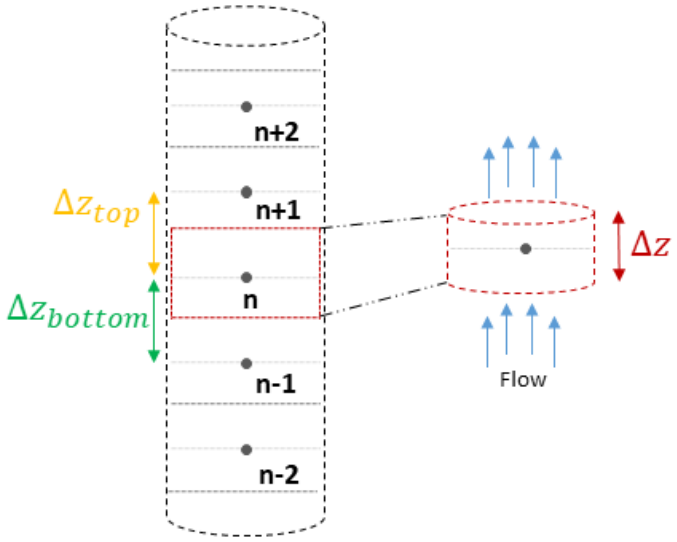

Figure 3.2: Finite Difference scheme for the wellbore discretization

Isolating the transient terms from equation (2-11):

$$\frac{\partial p^{wb}}{\partial t} - \frac{\beta_o}{c_o}\frac{\partial T^{wb}}{\partial t} = -\frac{Q}{A}\frac{\partial p^{wb}}{\partial z} + \frac{Q\beta_o}{Ac_o}\frac{\partial T^{wb}}{\partial z} - \frac{1}{Ac_o}\frac{\partial Q}{\partial z}. \tag{3-10}$$

According to Fig. 3.2, the central derivative of pressure and temperature can be expressed as:

$$\frac{\partial p^{wb}}{\partial z} = \frac{p^{wb}_{n+1} - p^{wb}_{n-1}}{\Delta Z_{top} + \Delta Z_{bottom}} \tag{3-11a}$$

$$\frac{\partial T^{wb}}{\partial z} = \frac{T^{wb}_{n+1} - T^{wb}_{n-1}}{\Delta Z_{top} + \Delta Z_{bottom}} \tag{3-11b}$$

The central derivative of the flow is represented by the following equation:

$$\frac{\partial Q}{\partial z} = \frac{Q_{n+1} - Q_{n-1}}{\Delta Z_{top} + \Delta Z_{bottom}}. \tag{3-12}$$

The flow rate present in the two first terms of the right hand side of Equation (3-10) is evaluated in the previous time step to linearize the resulting system of algebric equations. This linearization is being adopted since the flow quickly reaches a steady-state condition. The discretization of the mass balance equation is given by:

$$\begin{aligned}\frac{p_n^{wb,j+1}-p_n^{wb,j}}{\Delta t}-\left(\frac{\beta_o}{c_o}\right)\frac{T_n^{wb,j+1}-T_n^{wb,j}}{\Delta t} &= \theta[-\frac{Q_n^j}{A}\left(\frac{p_{n+1}^{wb,j+1}-p_{n-1}^{wb,j+1}}{\Delta Z_N+\Delta Z_S}\right)\\ &+\frac{Q_n^j\beta_o}{Ac_o}(\frac{T_{n+1}^{wb,j+1}-T_{n-1}^{wb,j+1}}{\Delta Z_N+\Delta Z_S})-\frac{1}{Ac_o}(\frac{Q_{n+1}^{j+1}-Q_{n-1}^{j+1}}{\Delta Z_N+\Delta Z_S})]\\ &(1-\theta)[-\frac{Q_n^j}{A}\left(\frac{p_{n+1}^{wb,j}-p_{n-1}^{wb,j}}{\Delta Z_N+\Delta Z_S}\right)+\frac{Q_n^j\beta_o}{Ac_o}(\frac{T_{n+1}^{wb,j}-T_{n-1}^{wb,j}}{\Delta Z_N+\Delta Z_S})\\ &-\frac{1}{Ac_o}(\frac{Q_{n+1}^{j}-Q_{n-1}^{j}}{\Delta Z_N+\Delta Z_S})].\end{aligned} \tag{3-13}$$

Discretizations of the energy conservation (Equation(2-21)) and momentum (Equation(2-28)) equations are presented in appendix B and C.

3.2 Initial and boundary conditions of the coupled wellbore-reservoir system

A set of initial and boundary conditions are defined to express the physical behavior of the system. Before the well top valve is opened, it is assumed the well is filled with oil. Therefore, initial pressure and temperature condition need to be in hydrostatic and geothermal equilibrium, respectively.

$$p^{wb}(z,t=0)=p^0-\rho g\sin(\alpha); \tag{3-14}$$

$$T^{wb}(z,t=0)=T^0-g_{G_z}\sin(\alpha); \tag{3-15}$$

$$Q^{wb}(z,t=0)=0; \tag{3-16}$$

$$p_r(r,t=0)=p^0; \tag{3-17}$$

$$T_r(r,t=0)=T^0; \tag{3-18}$$

Where p^0 and T^0 are the pressure and temperature reference at $z=0$ and $t=0$.

The boundary condition related to the flow rate is defined at the top of the well. During the drawdown period, the flow rate is set to a constant value

Q, and during the buildup period, the flow rate is set to zero, $Q = 0$. Pressure and temperature at the end of the reservoir ($r = r_e$) are assumed to be p^0 and T^0, respectively, and this condition is guaranteed by using a large value of r_e.

$$Q^{wb}(z = L, t > 0) = Q; \tag{3-19}$$

$$p_r(r = r_e, t > 0) = p^0; \tag{3-20}$$

$$T_r(r = r_e, t > 0) = T^0; \tag{3-21}$$

Mass balance and thermal energy balance are used in a control volume to evaluate the coupling condition between the well and reservoir. The mass balance without wellbore storage effect yield the following expression:

$$2\pi r_{co} h(\rho_r V_r) = \rho^{wb} Q^{wb}; \tag{3-22}$$

$$\frac{2\rho_r K}{(r_{co}/\mu)}\left(\frac{dp_r}{dr}\right) = \rho^{wb}\frac{Q^{wb}}{(\pi r_{co}^2 h)} \tag{3-23}$$

This equation contains a variable that refers to the well domain, Q^{wb}. Where the flow rate Q^{wb} entering the well is equal to the flow leaving the reservoir.

A simple thermal energy balance is assumed, leading to the following coupling boundary conditions:

$$P_{r_1} = P^{wb_1}; \tag{3-24}$$

$$T_{r_1} = T^{wb_1}; \tag{3-25}$$

The premise made in this model is that all the flow and energy transport that reache the well come from a source point, which leads us to consider that the first node of the reservoir coincides with the first node of the well.

To smooth out possible oscillations generated at the beginning of the

temperature and pressure derivatives and to get closer to reality, this work used an exponential function instead of a step function to control the flow.

$$Q(t) = Q_w(1 - exp(-t)), \tag{3-26}$$

where Q_w is the desired flow rate. Figure 3.3, has a comparison between the flow when considered as an exponential function ("Flow(Exp)") and the step function ("Flow(Step)") used by Da Silva[21] and Gonçalves [22].

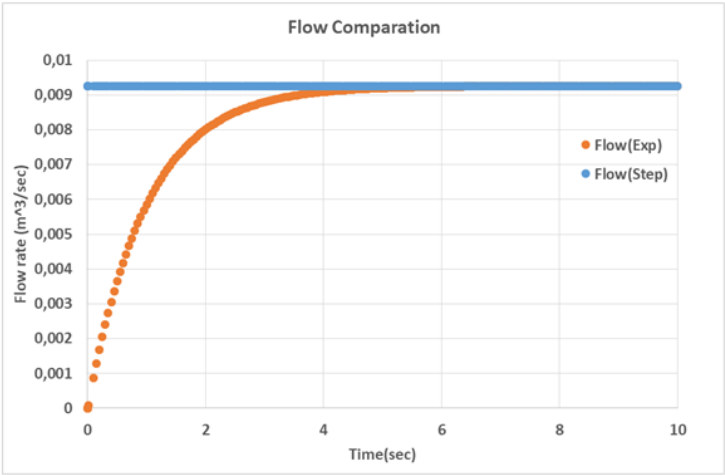

Figure 3.3: Comparison of flow when considered an exponential("Exp(Step)") function and a step ("Flow(Step)") function.

Figure 3.3 shows that the exponential equation reaches constant flow in less than 6 seconds, showing that it was only used to smooth the initial flow behavior.

Figure 3.4 presents a schematic representation of the reservoir and well discretization, illustrating the coupling between both domains.

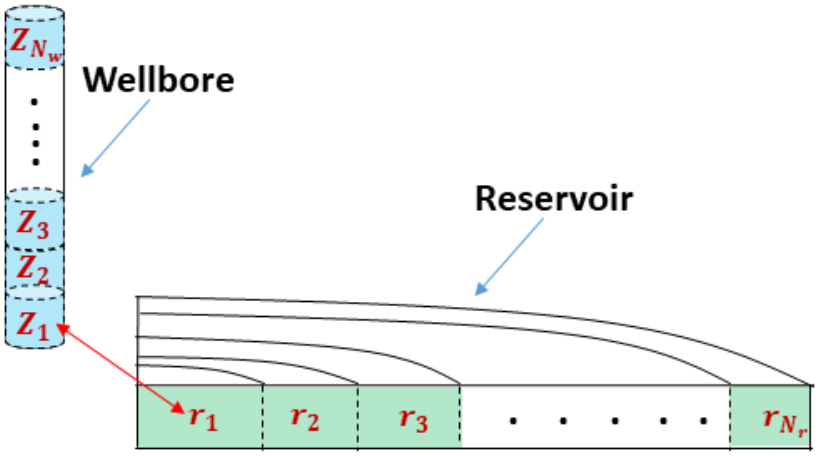

Figure 3.4: Scheme for the wellbore reservoir coupling

3.3 Mesh grid

Usually, the finite difference method requires a refined mesh, which leads to a high computational cost during transient problems simulations. To minimize computational cost without compromizing accuracy, it is usual to use a non-uniform mesh, where the nodes are concentrated in regions where large gradients of the variables are observed. Specifically, this work finds the region with higher pressure and temperature gradients in the reservoir near the coupling region with the wellbore.

A non-uniform reservoir grid was generated using a hyperbolic function presented by Vinokur (1983)[70], as shown in the following Equation:

$$r(i) = r_w + \left[1 + \frac{tanh(A[(i-1)/(N-1) - 1])}{tanh(A)}\right](r_e - r_w). \tag{3-27}$$

A represents the concentration factor parameter, and i is the node index, $1 \leq i \leq N_r$. The subscript w and e refer to the radius of the well, beginning of the reservoir, and the external radius of the reservoir, respectively. The well mesh was uniform.

The concentration factor (A), and also the number of nodes used to ensure that the results are independent of the mesh are shown in Chapter.4.

3.4 Time discretization

Establishing good criteria for the time step evolution is necessary to avoid numerical instabilities in the solution's problem. The set of equations that is

solved is non-linear. Some variables are evaluated in the previous step time using the fixed point iteration to linearize the equations.

The direct solver considers conveniently variable time steps by using a sigmoid function (SMF), defined as follow:

$$f(a,b,t) = \begin{cases} 2(\frac{t-a}{b-a})^2 \ , \ a \leq t \leq \frac{a+b}{2} \\ 1 - 2(\frac{t-a}{b-a})^2 \ , \ \frac{a+b}{2} \leq t \leq b \\ 1 \ , \ t \geq b \end{cases} \ , \tag{3-28}$$

Where a is the parameter determining the initial time step increment and b is the end time when the SMF function will be asymptotic. Then, the normalized value of the SMF function is multiplied by the maximum Δt (Δt_{max}), allowing to determine different times step, Δt.

Figure 3.5 shows an example of how each parameter of the SMF function influences the time step. For example, a time series from 0 to 100 with a step of 0.1 was created.

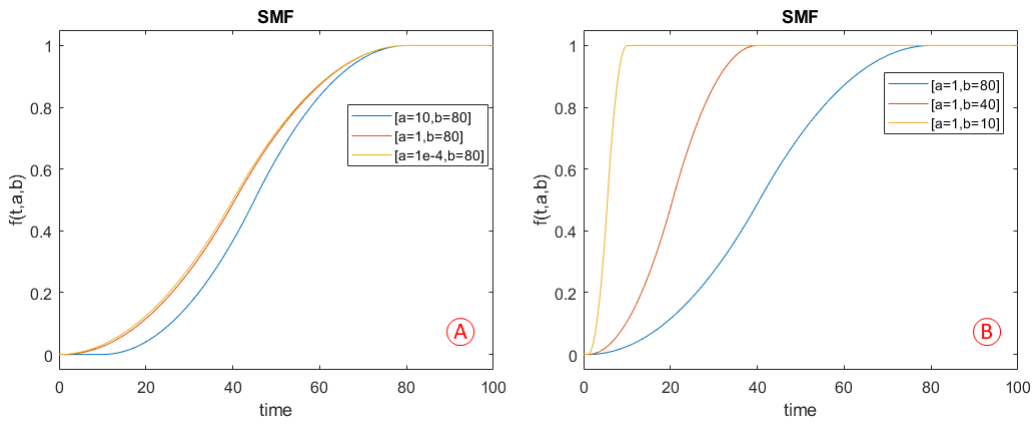

Figure 3.5: Example of the parameters influence into the SMF function.

Figure 3.5 A) shows the influence of the initial time a, in which values higher than 1 make a concentration of points with small Δt values. Figure 3.5 B) shows that parameter b from the SMF function is the point where the SMF achieves the maximum value and the Δt becomes $\Delta t_{max} = 60s$. This work set $a = 10^{-4}$ and $b = 80$ for a points distribution from 0 to 1200 with a step of 0.01.

3.5 Flowchart of the direct and inverse problem algorithms

The flow simulator solves a linear system at each instant of time, as illustrated in figure 3.6. This linear system involves solving three equations in each well node and two equations in each reservoir node. The simulator consists of a system of $(3N_w + 2N_r)$ equations, being solved at each time step of the simulation, where N is the number of nodes into the mesh domain (wellbore and reservoir).

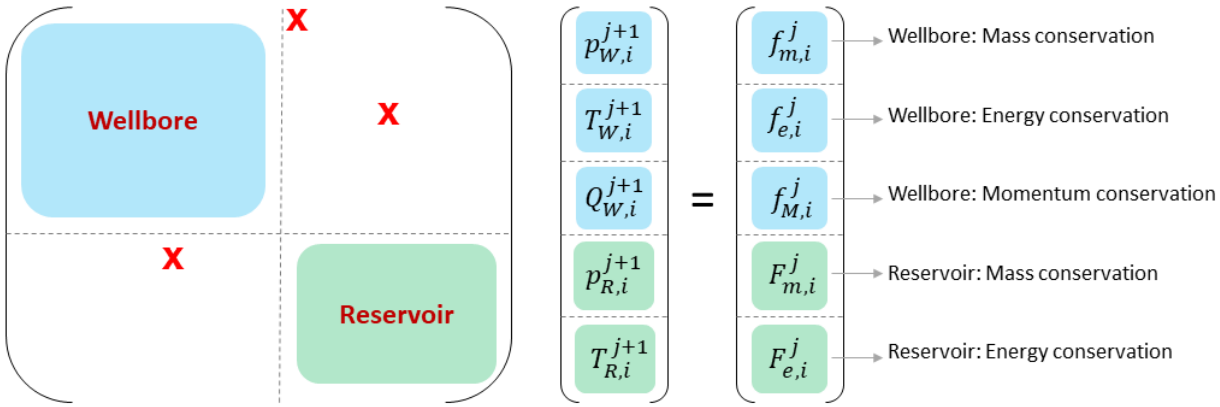

Figure 3.6: Scheme of the linear system that the simulator solves at each time step. The region in blue belongs to the wellbore domain, and in green is the domain associated with the reservoir. The red cross is the representation of the coupling comditions.

Figure 3.7 shows the flowchart of the inverse problem. Initially, the observed data is loaded, together with the input parameters of the inverse problem. The input consist of the $\mathbf{C}_D$ matrix, the number of assimilations(N_a) that will be made, the values of α_{factor}, and the ensemble of the vectors of parameters.

Then, the assimilation loop is started where $\mathbf{d}_{duc}$ value is calculated. The ensemble of vectors is passed to the simulator to run in parallel and produce the simulated data.

After that, the matrix $\mathbf{D}$ is created using the simulated data, and calculated the matrix $\mathbf{C}_{DD}$ and $\mathbf{C}_{mD}$. Those matrices are used to update the ensemble of vectors of parameters. This process is repeated until achieving the total number of assimilations (N_a)

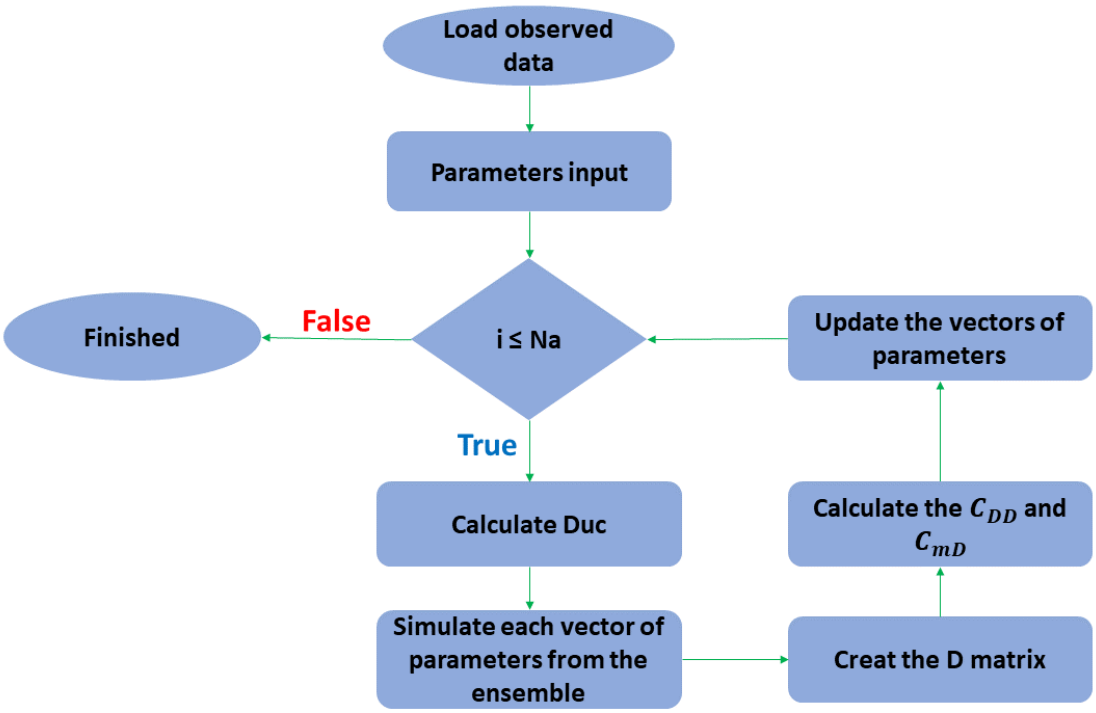

Figure 3.7: Flowchart of the inverse problem.

3.6 Computational cost

The simulations was performed into a computer with the following configurations:

- Windows 10 pro;
- Intel core i9-7920xCPU@2.90Ghz;
- 128Gb of ram memory.

The following table contains the time it took to run a simulation considering different amounts of simulated days, as well as the properties.

The following table contains the time it took to make the estimations considering the differents cases. The discription of each case is in the chapter 5.

Direct Problem		
Input properties	Simulated time (Days)	Simulation time (Seconds)
Galvão et al.[11]	2	54.12
Galvão et al.[11]	4	76.37
Onur et al.[5]	20	436.89

Table 3.1: Time required to run a simulation considering different amounts of simulated days, as well as properties.

Inverse Problem		
Case name	Simulated time (Days)	Simulation time (Seconds)
Case 1	2	2.9×10^3
Case 2	4	4.61×10^3
Case 3	4	6.16×10^3

Table 3.2: Time required to make the estimation considering different amounts of simulated days, as well as reservoir configurations.

4
validation of the direct problem solver

The implemented numerical code was validated using a commercial program (CMG-Stars) and with some analytical solutions presented in the literature. Initially, the validation process is done in the reservoir domain and then considering the coupled wellbore-reservoir system.

4.1
Reservoir domain validation

The reservoir model presented by Onur and Cinar (2016)[5] was used here for validation purposes. The model consists of a homogeneous cylindrical reservoir producing during 5 days at a constant flow of $Q_{sc} = 3.13\text{x}10^{-3}$std m^3/s (drawdown), followed by 15 days of static (buildup). Tables 4.1 and 4.2 show the rock and fluid properties used as input data. Results using the STARS-Thermal Advanced Processes Simulator from CMG, considering a grid with 200 blocks in the radial direction, were presented by Onur and Cinar (2016)[5].

K	Permeability[m^2]	1.056×10^{-13}
ϕ	Porosity	29%
T^o	Temperature[K]	351.48
p^o	Pressure[MPa]	13.06
H	Reservoir thickness[m]	30.48
r_w	Wellbore radius[m]	0.125
r_e	External Radius[m]	15000

Table 4.1: Rock and fluid properties extracted from Onur and Cinar (2016) [5], part-I.

C_o	Oil Compressibility [Pa^{-1}]	1.077×10^{-9}
C_w	Water Compressibility [Pa^{-1}]	4.398×10^{-10}
C_r	Rock Compressibility [Pa^{-1}]	4.351×10^{-10}
C_t	Total Compressibility [Pa^{-1}]	1.417×10^{-9}
c_{po}	Oil Heat Capacity [J/kgK]	2177.1
c_{pw}	Water Heat Capacity [J/kgK]	4186.8
c_{pr}	Rock heat capacity [J/kgK]	962.96
s_w	Water saturation	0.15
B_o	Oil formation volume factor [m^3/std m^3]	1.05427
λ_t	Thermal Conduct. Porous medium [J/msK]	3.4615
ρ_w	Water density [kg/m^3]	1000.03
ρ_r	Rock density [kg/m^3]	2643.05
ρ_o	Oil density [kg/m^3]	834.56
β_o	Oil thermal-expansion coefficiet[K^{-1}]	7.2×10^{-4}
β_r	Rock thermal-expansion coefficiet[K^{-1}]	9.0×10^{-5}
β_w	Water thermal-expansion coefficiet[K^{-1}]	9.0×10^{-4}
μ_o	Oil viscosity [Pa.s]	2.949×10^{-3}
α_t	Thermal diffusivity total [m^2/s]	1.42×10^{-6}
φ^*	Effective adiabatic expansion coefficient[K/Pa]	2.31×10^{-8}
φ_w	Water adiabatic expansion coefficient[K/Pa]	4.554×10^{-8}
φ_o	Oil adiabatic expansion coefficient[K/Pa]	1.72×10^{-7}
ε_{JTo}	Joule-Thomson expansion coefficient[K/Pa]	-4.432×10^{-7}
$(\rho C_p \varphi)_t$	Adiabatic-expansion coefficient of the fluid system[$J/m^3 Pa$]	5.63×10^{-2}
$(\rho C_p)_t$	Volumetric heat capacity of the fluid-saturated rock[$J/m^3 K$]	2.437×10^6

Table 4.2: Rock and fluid properties extracted from Onur and Cinar (2016) [5], part-II.

Figure 4.1 shows transient pressure data as a result of 20 days of reservoir test with 15 days of buildup. The mesh test was also performed, concluding that 50 nodes and concentration factor $A = 8.5$ are enough to reproduce the results presented in the literature.

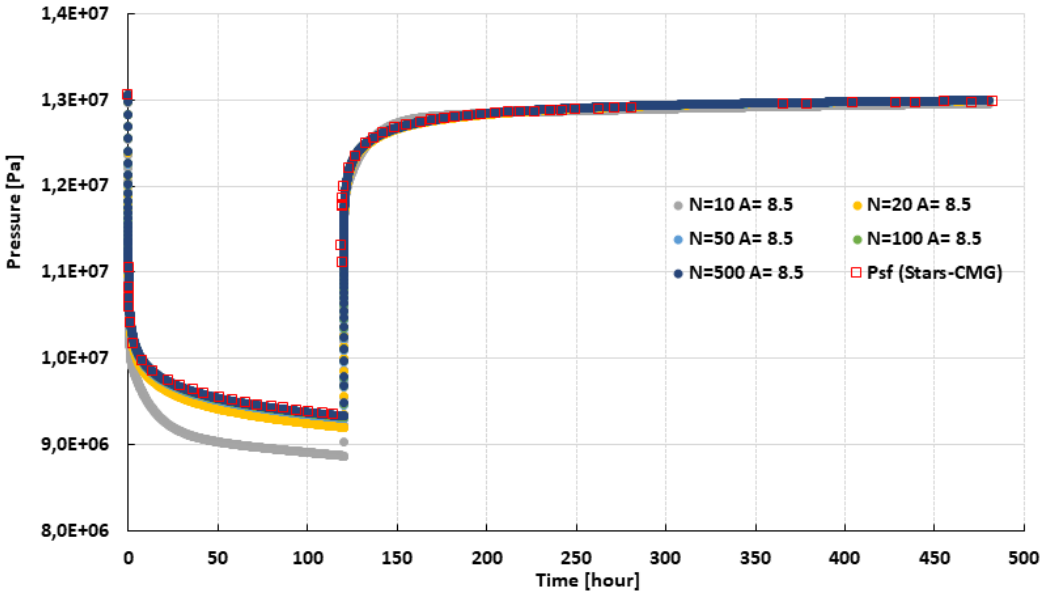

Figure 4.1: A comparison plot between the pressure evolution presented by Onur and Cinar (2016) and the evolutions from the flow simulator for different amounts of nodes in the reservoir grid.

Transient temperature data obtained using the simulator also need to be validated. Figure 4.2-A contains mesh test results in the semi-log scale during the drawdown period. Here, it also is possible to observe that 50 nodes with $A = 8.5$ are enough to ensure the temperature data validation. Figure 4.2-B contains mesh test results in the semi-log scale during the buildup period. Where also, 50 nodes with $A = 8.5$ are enough to ensure the temperature data validation.

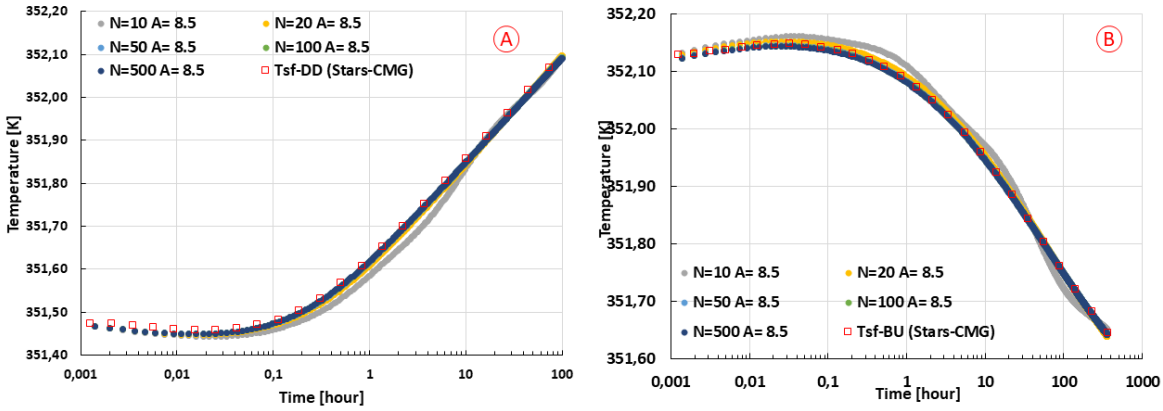

Figure 4.2: Plot A is the comparison between the evolution of the semi-log temperature presented by Onur and Cinar (2016)[5] and the evolutions coming from the flow simulator for different amounts of nodes in the reservoir grid during the flow period. The plot B is the comparison durign the static period (buildup).

After a mesh test and validation, it is possible to conclude that 50 nodes along the reservoir domain are enough to obtain mesh-independent solutions. It is vital to notice that the validation is being applied in a homogeneous medium where the permeability and porosity are constant. However, a heterogeneous reservoir will be considered in this work, then 300 nodes with a concentration factor of A=8.5 is used to ensure mesh independent solutions.

4.1.1 Fully coupled well-reservoir validation

Results obtained by Galvão et al. [71] were used to validate the transient data along the well coupled to a reservoir. Galvão et al. [71] validated their analytical solution of a homogeneous reservoir comparing with results obtained using CMG-Stars. The coupling process in this simulator consists of a two-dimensional cylindrical reservoir, in which the mesh has 200 grids in the radial direction and 41 layers in the vertical direction. However, the first 40 layers are not producing, representing the formation adjacent to the well, so the model represents a one-layer reservoir coupled to a well. Properties of the reservoir and the well are shown in tables 4.3 and 4.4.

For this validation, a sequence with 24 hours of the drawdown at a constant flow rate of $1400 m^3/day$ followed by 48 hours of buildup is considered. Figure 4.3 shows the comparison between the temperature of the bottom hole (yellow triangles) and the wellbore head (red cross and blue circles). Due to the full open hypotheses, the bottom hole pressure and temperature are equal

to the sandface pressure and temperature. The temperature at the wellhead resulting from the solution of the direct problem is in blue, and the temperature obtained by the CMG-Stars is in red.

K	Permeability[m^2]	9.87×10^{-14}
ϕ	Porosity	12%
T^o	Temperature[K]	334.0
p^o	Pressure[MPa]	49.033
H	Reservoir thickness[m]	50
r_w	Wellbore radius[m]	0.156
r_e	External Radius[m]	25000
C_o	Oil Compressibility [Pa^{-1}]	1.12×10^{-9}
C_w	Water Compressibility [Pa^{-1}]	4.04×10^{-10}
C_r	Rock Compressibility [Pa^{-1}]	3.06×10^{-10}
c_{po}	Oil Heat Capacity [J/kgK]	2252.9
c_{pw}	Water Heat Capacity [J/kgK]	4209.35
c_{pr}	Rock heat capacity [J/kgK]	888
B_o	Oil formation volume factor [m^3/std m^3]	1.4
s_w	Water saturation	0.15
λ_t	Thermal Conduct. Porous medium [J/msK]	3.44
ρ_w	Water density [kg/m^3]	998.2
ρ_r	Rock density [kg/m^3]	2643.05
ρ_o	Oil density [kg/m^3]	770.0
β_o	Oil thermal-expansion coefficiet[K^{-1}]	1.11×10^{-3}
β_r	Rock thermal-expansion coefficiet[K^{-1}]	9.0×10^{-5}
β_w	Water thermal-expansion coefficiet[K^{-1}]	5.27×10^{-4}
μ_o	Oil viscosity [Pa.s]	0.9×10^{-3}
α_t	Thermal diffusivity total [m^2/s]	1.484×10^{-6}

Table 4.3: Input data for the reservoir used to validate the direct problem solver and also used in Galvão et al. [11, 12]

L	Wellbore length[m]	512.5
r_{co}	Chasing external radius [m]	0.12224
r_{ci}	Chasing initial radius [m]	0.10839
r_{to}	Tubbing external radius [m]	0.06985
r_{ti}	Tubbing initial radius [m]	0.05931
$z_{tubbing}$	Tubbing position[m]	100.0
z_{gauge}	Gauge position[m]	512.5
λ_{cement}	Thermal Conduct. Cement [W/mK]	1.898
λ_{wall}	Thermal Conduct. Wall [W/mK]	44.917
λ_{an}	Thermal Conduct. Anulus [W/mK]	0.162
λ_{rock}	Thermal Conduct. Rock [W/mK]	3.8773
g_G	Geothermal gradient[K/m]	-0.03

Table 4.4: Input data of well parameters.

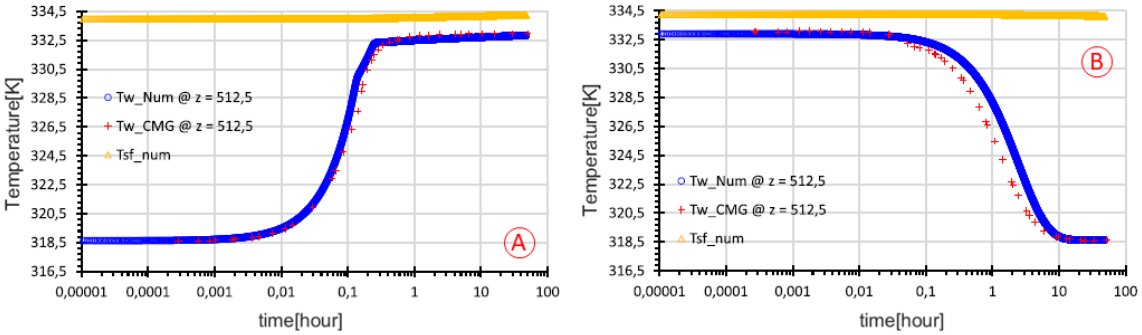

Figure 4.3: Plot (A) is the comparison between the evolution of the semi-log well head temperature using CMG (red cross), the evolutions from the flow simulator (blue circle), and the sandface temperature from the flow simulator (yellow triangle), during the flow period. Plot (B) is the same comparison but considering the static period.

In the well, at gauge position of $z = 512.5m$, the numerical results show effects of change in pipe area yielding two different flowing velocities at the same volumetric rate. The heating profile during the drawdown period, see Fig. 4.3 a), show the first change at approximately 0.15 h and it is caused by the elevation of the fluid originally at the bottom of the tubbing. The second change that occurs approximately 0.28 h is caused by the elevation of the fluid originally at the sandface, as discussed by Galvao et al.[11]. Showing a good agreement with the commercial simulator. During the buildup period, good agreement with the commercial thermal simulator is also observed, as shown in Fig. 4.3 b), with a slight difference in the interval of 0.02 h and 10 h, and above this period a good agreement is observed. Here, a uniform mesh with

300 nodes along the well is used.

COLOCAR AQUI O TEMPO PARA CADA ANÁLISE DO PROBLEMA DIRETO!

5
Results of the inverse problem

This chapter initially describes the three different reservoir configurations where the ES-MDA method is applied to estimate different reservoir parameters. After that, a sensitivity analysis of the $\mathbf{C}_D$ matrix from equation 2-66 is discussed, considering both pressure and temperature with the first reservoir configuration as observed data . The ES-MDA method is then used with the two most accurate $\mathbf{C}_D$ matrix configurations from the previous analysis in the other reservoir configurations. Estimations made considering both pressure and temperature data are compared with the estimations made with only pressure data. In all examples, the observed data is a synthetic data obtained by adding a Gaussian noise in the solution of the direct problem for a specified set of parameters. Finally, a robustness test of the ES-MDA method is done by adding periodic disturbances (such as tidal effect) in the pressure observed data with different amplitudes.

5.1
Reservoir configurations considered in the analyses

In this work, three different reservoir configurations are analyzed. The reservoirs are cylindrical and one-dimensional, with properties equal to cases presented by Galvao et al. [11].

Case-1 represents a reservoir with a damage zone (skin) near the well (skin radius $0.78m$) with a permeability of $32mD$ followed by a homogeneous region with a permeability of $100mD$. Porosity is assumed to be constant and equal to 0.12. Here, the parameters to be estimated are both permeabilities, the skin radius, and the porosity.

Case-2 has three different permeability zones. Here, the damage zone near the well has $500mD$ of permeability with a known skin radius of $1m$. A transition zone is created with a permeability of $200mD$ extended to a radius of $2.5m$. A homogenous region with $100mD$ represents the final part of the reservoir. The porosity is also assumed to be constant to 0.12 along with the all reservoir domain. Here, the parameters to be estimated are three permeability regions, the radius of the transition zone, and the porosity.

Case-3 has similar configurations to Case-2. The only difference between

them is the extended radius of $100m$ of the transitions zone. In this case, the parameters to be estimated are six: the permeabilities of three different regions, the skin radius, the radius of the transition zone, and the porosity.

Figure 5.1 represents the three different cases with all dimensions and properties that will be analyzed in this work. The variables in red represent the parameters to be estimated.

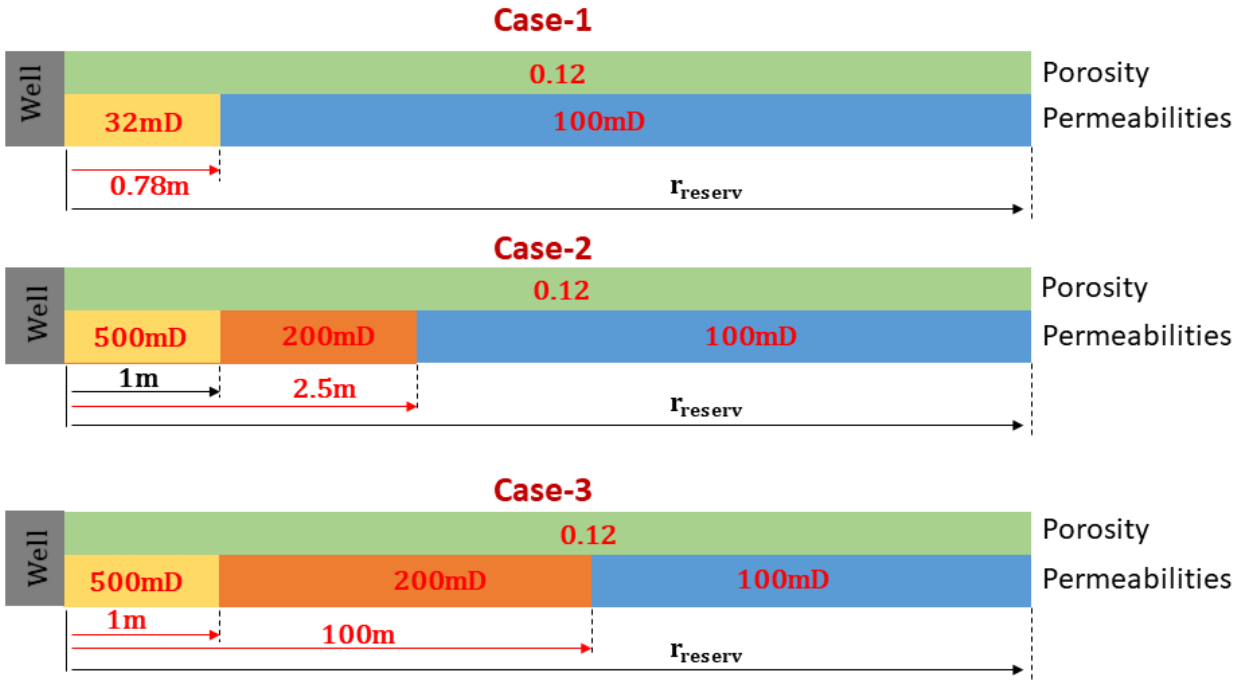

Figure 5.1: Scheme of the three cases that will be analyzed in this work. Each case has the properties values and zone dimensions, the values in red are the values of the variables in the estimation process.

5.2 Sensibility analyses of $\mathbf{C}_D$ matrix

In this section, the effect of the $\mathbf{C}_D$ matrix, used in equation (2-66) in the parameter estimation is analyzed. The analysis was motivated by the different pressure and temperature data scales that could lead to inefective updates. Another motivation for this analysis is a lack of information on temperature data coupled with pressure data applied to the ES-MDA method in the literature.

As mentioned in Chapter 2, the $\mathbf{C}_D$ matrix is always diagonal, and the analysis here consists of finding appropriate coefficients for pressure and temperature data.

Firstly, the transient pressure and temperature data are normalized using their respective initial values (p^o and T^o).

The $\mathbf{C}_D$ matrix comprises two diagonal blocks associated with each independent observed data. The first block corresponds to pressure data, as shown in Fig. 5.2.

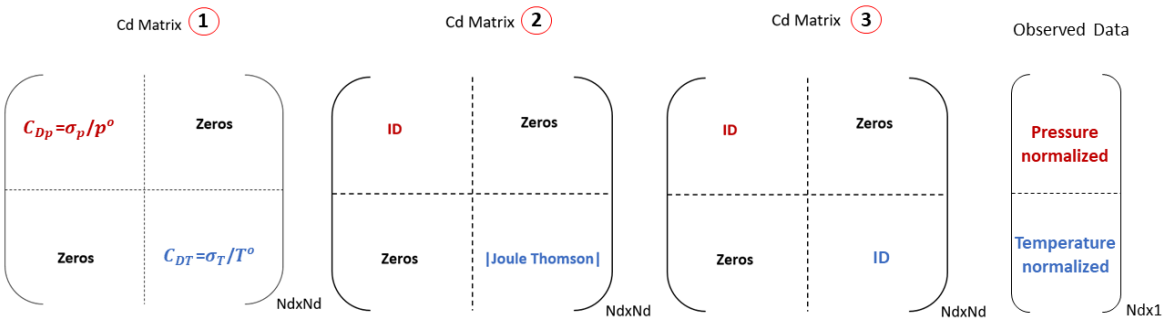

Figure 5.2: Schematic of the $\mathbf{C}_D$ matrix configurations used and the couple observed data.

Figure 5.2 shows the configurations of the $\mathbf{C}_D$ matrix considered. The first one is based on Silva [4] and Xu and Forouzanfar[26]. A standard deviation is assumed for both pressure and temperature data, $\sigma_p = 10$Bar[4] and $\sigma_T = 0.005$K[26]. Since the observed data is normalized, the standard deviation values used in $\mathbf{C}_D$ found in the literature (σ_p and σ_T) are also normalized by their respective initial values (p^o and T^o). This normalization procedure was performed to reduce the difference between the data magnitude and facilitate the procedure to invert the matrix. In this work, the inverse of the matrix was made using the command $inv(matrix)$ from Matlab.

Therefore the first block of $\mathbf{C}_D$ matrix is a diagonal with $\mathbf{C}_{Dp} = 2.04 \times 10^{-5}$, and the other block is $\mathbf{C}_{DT} = 7.48 \times 10^{-12}$. The values of the second block are in the order of 10^7 smaller than the value of the first block. This proportion is in the same order as the magnitude of the Joule Thomson coefficient for the oil ($|\varepsilon_{JTo}| = 3.44 \times 10^{-7}$).

The second configuration of $\mathbf{C}_D$ matrix was based on Sui [46], at which the observation weight matrix is assumed 1 for pressure data and the Joule Thomson coefficient (ε_{JTo}) for the temperature data. This configuration used by Sui was applied in the Levenberg-Marquardt method

The third configuration for the $\mathbf{C}_D$ matrix is an identity matrix. Four alternatives configurations are considered, by simply multiplying the first and the second configurations by a constant factor(weight), equal to 10^{-2} and 10^{-4}.

In order to facilitate the notation, the first configuration from now on is called Cdp/Cdt, the second is $ID/EJTO$ and the identity configuration

is ID. The followed number in the notation refers to the weight value. So $(Cdp/Cdt)(1e-2)$ means the first $\mathbf{C}_D$ matrix configuration considering the weight of 1.0×10^{-2}.

Comparison between of the different $\mathbf{C}_D$ matrix configurations was made by comparing the solution of the inverse problem for Case-1. For this analysis, the ensemble size was fixed at 100 (N_e=100), the inflation factor (α_{factor}) was taken constant equal to 4, and the ES-MDA performed four assimilations.

A common practice to estimate permeabilities values in the inverse problem context is using a log scale. Therefore, the initial ensemble for permeabilities in terms of log scale is created with a normal distribution with a mean equal to 5.0 and standard deviation equal to 1.0 ($N(5.0, 1.0^2)$), as used by Emerick 2012 [54].

The skin radius (r_{skin}) is estimated as a multiple of the well radius (r_w), named as α_{skin} such that:

$$r_{skin} = \alpha_{skin} \times \ r_w \tag{5-1}$$

For the initial α_{skin} ensemble, a normal distribution with a mean equal to 5.0 and a standard deviation equal to 1.0 ($N(5.0, 1.0^2)$) is also used. For the porosity, the normal distribution had a mean of 0.2 and a standard deviation of 0.05 ($N(0.2, 0.5^2)$)[54].

The vector of parameters $\mathbf{m}$ that this comparison is trying to estimate is represented by equation 5-2:

$$\mathbf{m} = [log(K_1),\ log(K_2),\ \alpha_{skin}\ and\ \phi]^T. \tag{5-2}$$

K_1 is the permeability of the skin zone, K_2 is the permeability out of the skin zone, α_{skin} is the radius of the skin zone, and ϕ is the porosity of the reservoir.

Figure 5.3 shows the initial distribution for the four unknown parameters for Case-1. Reverting the log permeability to millidarcy, this distribution ranged from $7.4mD$ to $1.1D$.

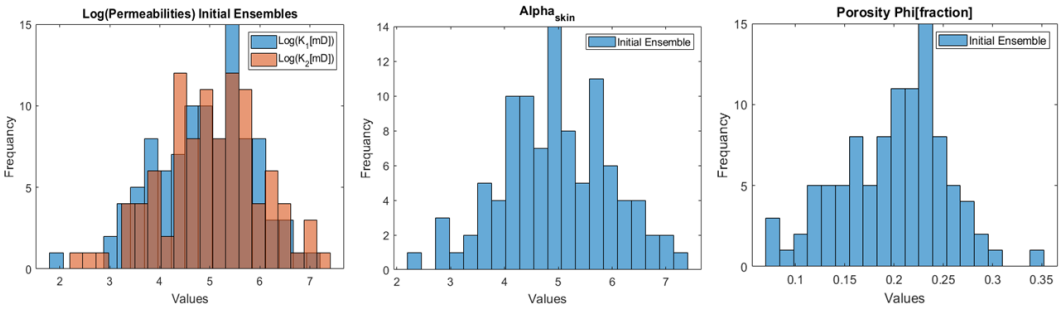

Figure 5.3: Initial distributions for the variabels in analysis into the Case-1.

For the $\mathbf{C}_D$ matrix analysis, the observed data was generated using the response of the flow simulator considering two days with a constant flow rate of $800m^3/day$. White noise is included trying to make the observed data more realistic.

$$d_{dobs} = Simulated\ Data + White\ Noise \tag{5-3}$$

The white noise is created considering a normal distribution with a zero mean for pressure and temperature data with a standard deviation of $50KPa$ $(0.5Bar)$ for pressure data and $0.005K$ for the temperature. Figure 5.4 shows the impact of this noise addition on both simulated data.

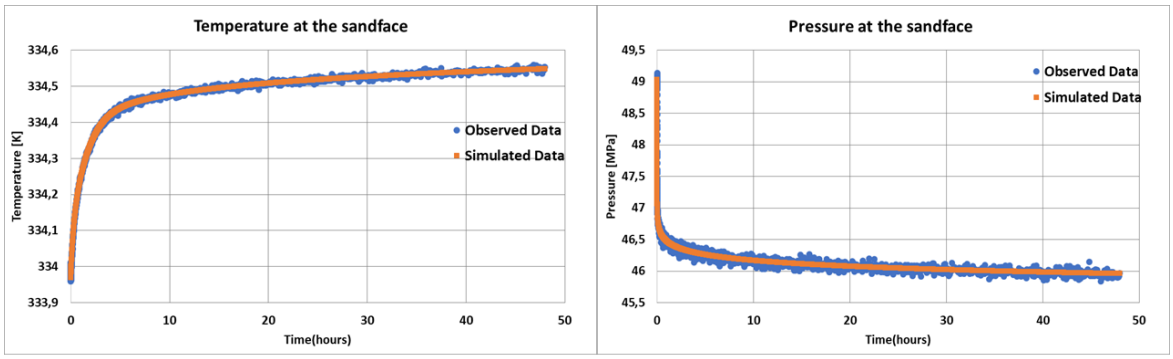

Figure 5.4: Comparison of temperature and pressure evolutions for the simulated data in orange and the observed data cosidering a white noise in blue.

Figures 5.6 to 5.11 presents the boxplots for the estimation of each parameter for Case-1. Figure 5.5 explains the interpretation of boxplots, which is a graphical method to represent the data through their quartiles. 50% of the data are inside the blue box and the red line, inside of the box, represents the median of the data. The points over the maximum and under the minimum

are considered outliers. The red line with the circles represents the expected value.

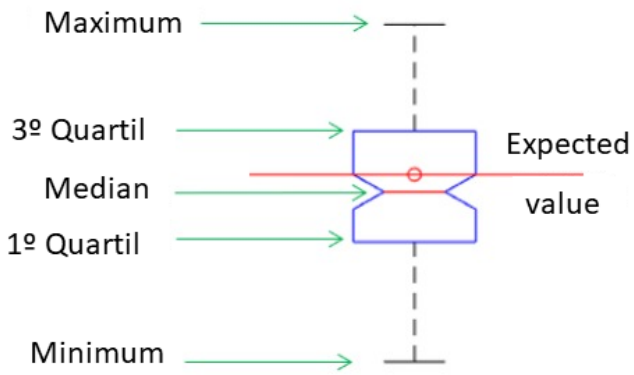

Figure 5.5: Comparison of temperature and pressure evolutions for the simulated Data in orange and the observed data cosidering a white noise in blue.

Figure 5.6 shows the estimation of the first permeability (K_1 or skin region). It corresponds to the first element of the vector of parameters ($\mathbf{m}$).

When the $\mathbf{C}_D$ matrix is considered an identity matrix, the ES-MDA method is not able to estimate the values close to the expected results and it does not reduce the uncertainty of the permeability K_1. The other tested configurations made estimations very close to the expected results (red reference line).

When we assume the $\mathbf{C}_D$ matrix as an identity matrix, we are considering that the data have low reliability, that is, they have large errors associated with their measurements, which results in a small reduction of uncertainties, as shown in the following graph.

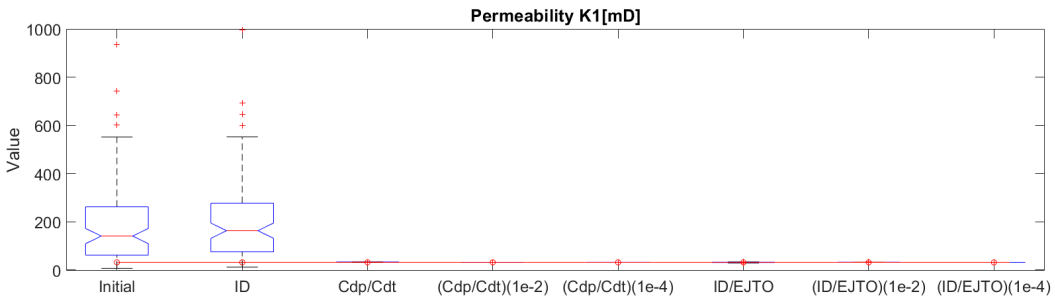

Figure 5.6: Boxplots for the skin permeability considering the initial distribution and all configurations of $\mathbf{C}_D$ matrix tested.

Figure 5.7 shows a zoom of Fig.5.6 for different $\mathbf{C}_D$ configurations, except ID. The line in red represents the expected value of K_1 to be achieved. From

this result, it is possible to observe that the ES-MDA method reduces the first parameter uncertainties that started with an initial distribution from $7.4mD$ to $1.1D$. According to the K_1 estimation, using $(ID/EJTO)(1e-4)$ and $(Cdp/Cdt)(1e-4)$ we obtained the best estimations.

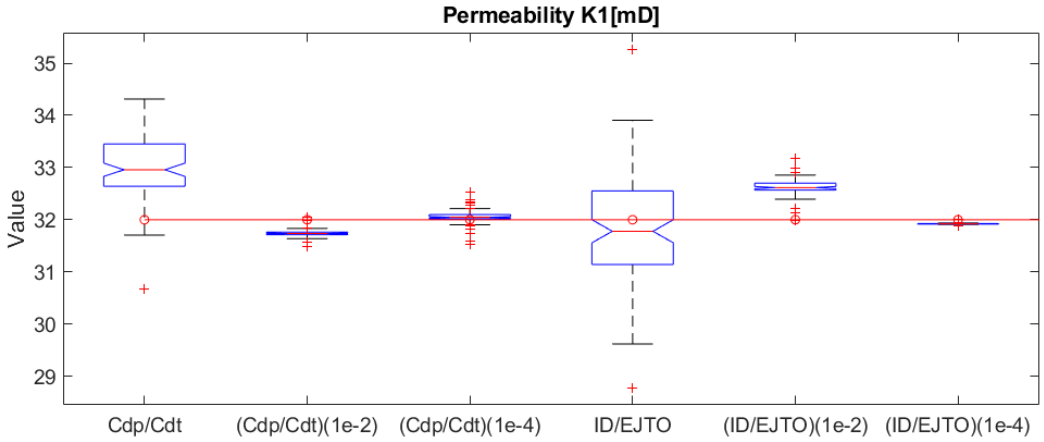

Figure 5.7: Zoom in the boxplot for the first permeability considering the best settings of $\mathbf{C}_D$ matrix.

Using the identity matrix again to estimate the second parameter (K_2), there is no regularization using the ES-MDA method, as shown in Figure 5.8. It happens in all parameters, and that is why results with identity matrix will not be shown for the following parameters. It is essential to mention that all four parameters are being estimated simultaneously here.

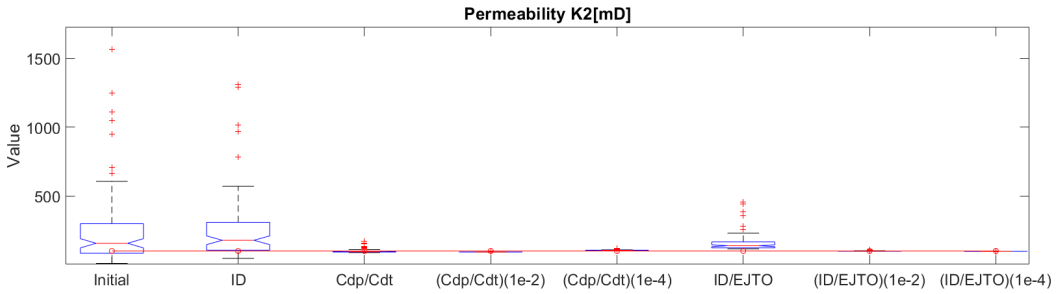

Figure 5.8: Boxplots for the second permeability, with the initial distribution and all configurations of $\mathbf{C}_D$ matrix tested.

Figure 5.9 shows a zoom of Fig. 5.8, where the red line represents the expected value of the second parameter (K_2), $100mD$. Comparing the estimations made using $ID/EJTO$ and (Cdp/Cdt) without weights with the

same configurations with the addition of weights, we observe that adding weights on the $\mathbf{C}_D$ matrix improved the parameter estimations.

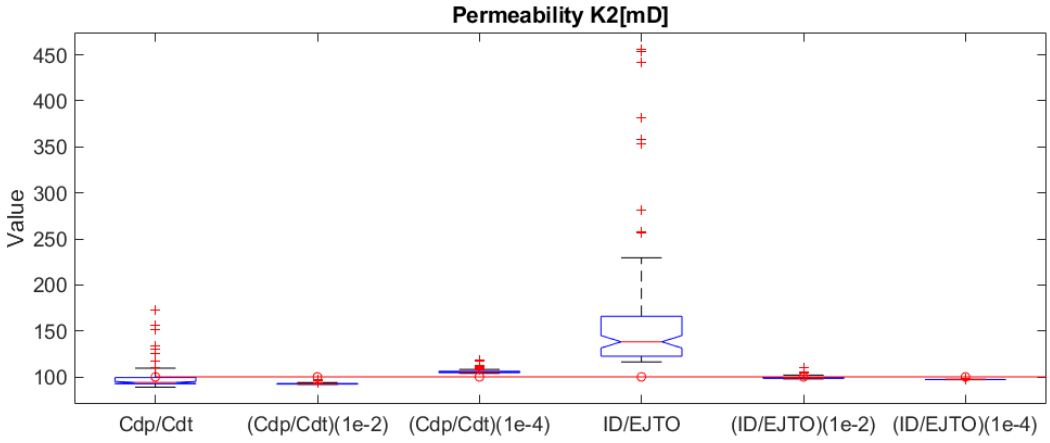

Figure 5.9: Zoom in the boxplot for the second permeability, considering the best settings of $\mathbf{C}_D$ matrix.

Figure 5.10 shows the results of the third parameter estimation, which represents the skin radius, α_{skin}. Once again, is evident that the addition of weight in the $\mathbf{C}_D$ matrix, improves the accuracy of the parameter estimative and reduces the uncertainty.

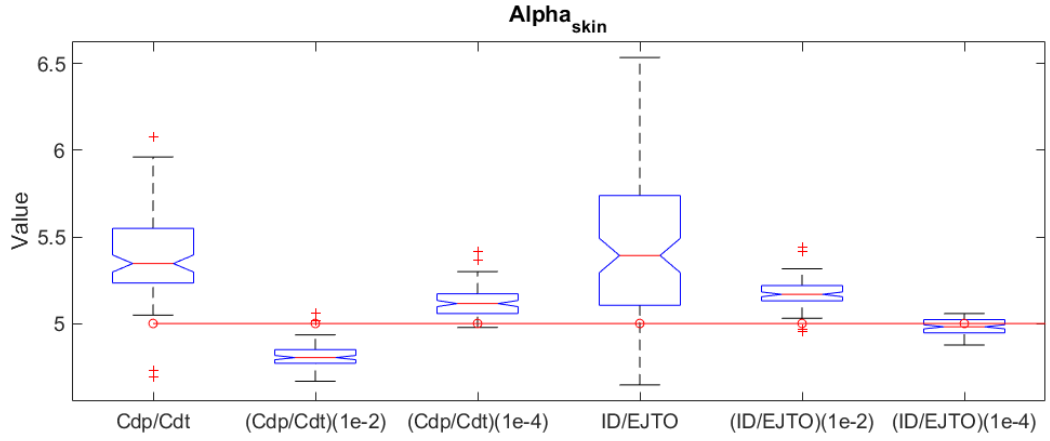

Figure 5.10: Boxplots for the α_{skin}, considering the best settings of $\mathbf{C}_D$ matrix tested.

Finally, Fig.5.11 shows the estimation of the porosity, which is the last element of the vector $\mathbf{m}$. Results with a higher degree of uncertainty are obtained when $ID/EJTO$ and (Cdp/Cdt) are used. However, applying appropriate weight values on the $\mathbf{C}_D$ matrix can decrease the level of uncertainty. In

this case, considering all of the parameter estimations, a factor of 1.0×10^{-4} produces an accurate estimation for the entire vector $\mathbf{m}$.

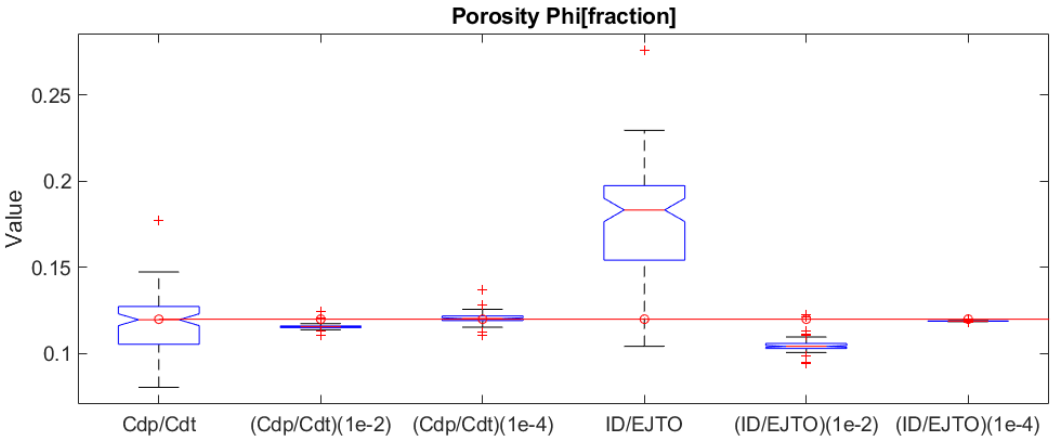

Figure 5.11: Boxplots for the porosity, considering the best settings of $\mathbf{C}_D$ matrix.

Based on the results shown in this section, we can conclude that considering the $\mathbf{C}_D$ matrix taken as an identity matrix does not generate almost any regularization of the analyzed parameters. It is also observed that there is an inverse relationship between the weight factor in the $\mathbf{C}_D$ matrix and the accuracy of parameter estimation.

The ES-MDA method keeping the best $\mathbf{C}_D$ matrix configuration from previous analysis is applied for those three reservoir configurations and compared with the estimative analysis made with only pressure data.

For the following analyses, we keep $(ID/EJTO)(1e-4)$ and $(Cdp/Cdt)(1e-4)$ as the $\mathbf{C}_D$ matrix configuration. In the next analyses, we compare the estimations made with only pressure data and the estimations made with both pressure and temperature data for the three reservoir configurations presented in Fig.5.1.

5.3 Case-1

To begin the analysis of this case, we present plots that are frequently used to make a diagnosis of the reservoir under study. The data analyzed in this section are generated by considering two days of constant flow $(800m^3/day)$.

The first analysis is based on the pressure data alone, since this approach has been used for a longer time and has a vast literature. Figure 5.12 is the composition of two plots. The right side was taken from the book (*Introduction*

to Well Testing) by Bath in 1998[6], which shows the different behaviors that the bottom hole pressure semi-log can present in different scenarios.

The plot on the left side of Fig. 5.12, shows the sandface pressure semi-log plot for the simulated data, which is the response of the direct problem, and observed data from Case-1.

Using the plot on the right as a reference, we can infer that the reservoir in analysis is similar to the case that has the skin zone with the addition of the storage effect. Thus generating a need for a deeper analysis for a better characterization. Based on the slope identified by the red line in the left plot, it is possible to estimate the reservoir's transmissibility value (Kh$\sim$ Permeability $\times$ reservoir height).

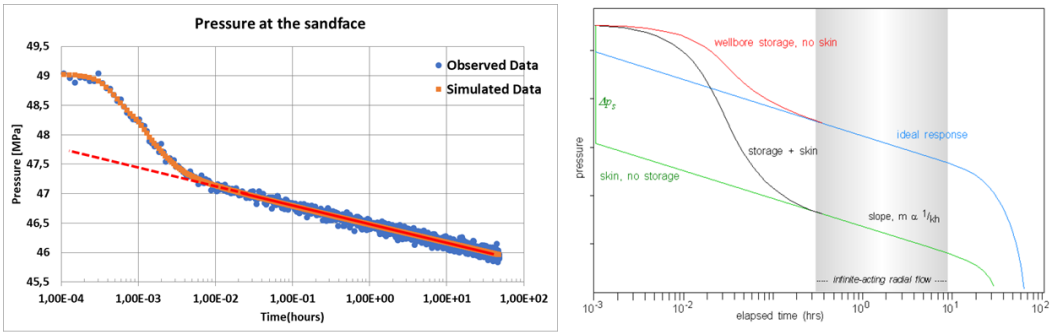

Figure 5.12: In the left is the semilog of the observed (in blue) and simulated (in orange) data. On the right is the bootom hole pressure semilog analysis of differents scenarios from the book Introduction to well testing by Bath[6].

Figure 5.13 presents the pressure derivative according to Bourdet et al. [72]. The right plot is an adaptation from the book (*Applied Well Test Interpretation*) from Spivey and Lee [7], in which the blue curve indicates the delta pressure ($p(\Delta t) - p(\Delta t = 0)$) and in red is the pressure derivative.

The single horizontal curve level that predominates in the pressure derivative curves, indicates radial flow in a homogenous reservoir. On the left side, is the derivative curves from the simulated and observed data. Due to the addition of the white noise into the observed data, the derivative pressure curve presented a greater degree of scattering of its points, making the identification of its levels somewhat difficult.

Analyzing the pressure derivative curve from the simulated data, we can infer that the pressure data indicates a reservoir that has a homogeneous permeability. This behavior is explained by the speed of propagation of the pressure diffusivity, which is so fast that the small region with different permeability near the well passes unnoticed by the pressure derivative data.

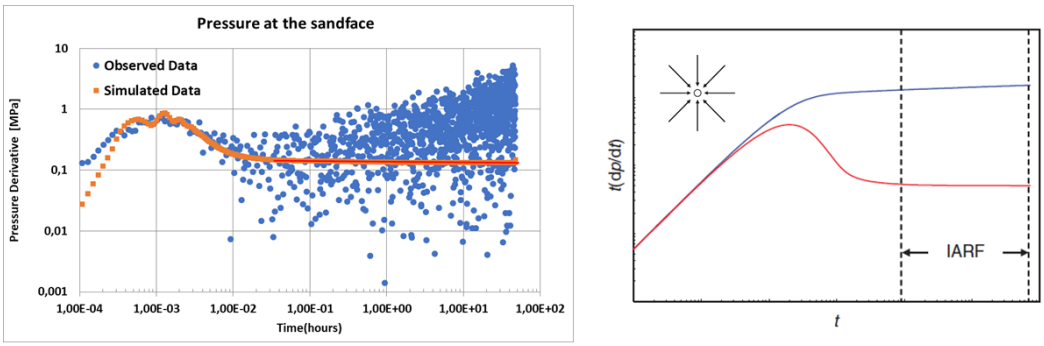

Figure 5.13: The graph on the left is the Bourdet derivative considering the simulated pressure data (in orange) and the observed data (in blue) which are the data simulated with the addition of white noise. The graph on the right was adapted from the book Applied Well Test Interpretation from Spivey and Lee [7], which represents the derivative curves from a homogeneous reservoir in a radial flow.

Due to the slower propagation of the temperature signal compared to the pressure signal, the temperature signal has been intensively studied to better characterize the region near the well. Compared to the literature on pressure data, the literature for temperature data is much smaller, in which Professor Onur([14],[5],[8],[3],[31]) has great relevance.

Figure 5.14 presents the semilog analysis from the sandface temperature. The plot on the right shows the results from a reservoir with skin (Synthetic Test 2) and without skin (Synthetic Test 1), presented by Onur et al.[8]. The slope of the curve indicates the variation of the transmissibility of a region in the reservoir. And as the reservoir model has a constant height, the changes in slope are due to changes in the permeability of the reservoir.

The left plot of Fig. 5.14 presents the semilog temperature data from the simulated and the observed data from Case-1 during the flow period. Different from pressure data, the temperature data can identify the different permeability regions near the well. These two different regions are identified by the slopes marked by the red lines, and by observing the time of this transition, we can infer that due to the lower velocity of the thermal propagation, the near well region can be better characterized using the temperature data.

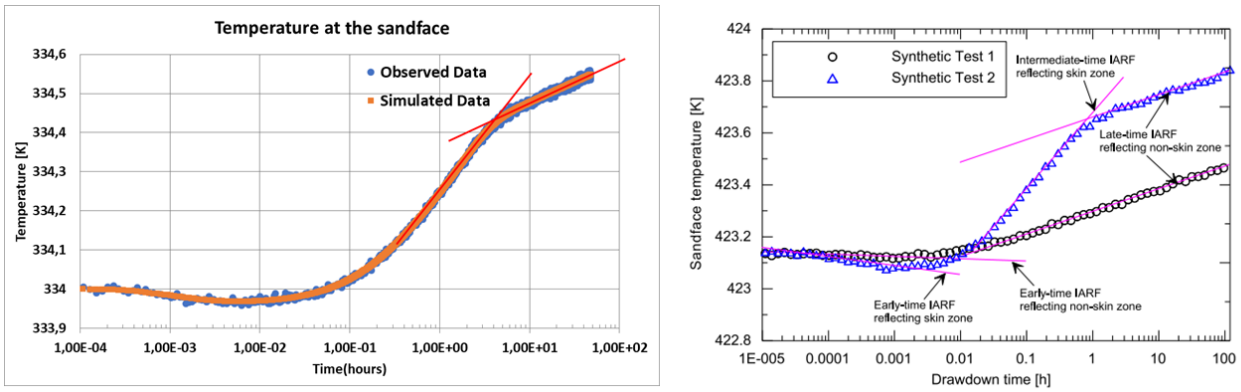

Figure 5.14: In the left is the semilog of the observed (in blue) and simulated (in orange) data. On the right is the semilog analysis of two synthetic reservoir presented by Onur et al.[8], which indicates that differents slopes into the semilog temperature analysis is referent of differentes permeability regions.

Figure 5.15 show the Bourdet [72] temperature derivative function, considering the temperature data presented by Onur et al.[8]:

$$Temperature\ Derivative = abs(T'_{ln\Delta t}) = \left|\frac{dT}{dln\Delta t}\right| \tag{5-4}$$

The plot on the right side of Fig. 5.15 was taken from Onur's works [8], in which the pink horizontal lines indicate permeability changes during the flow periods (Draw Down-DD) or static periods (Build Up-BU) in different reservoir configurations. The left plot shows the temperature derivative for the simulated and observed data during the DD.

As observed for the pressure data, the addition of white noise causes a dispersion of the temperature derivative, thus making it difficult to identify the levels reached by the temperature derivative.

Similar results presented by Onur related to temperature derivative data, we can also observe the different levels that the temperature derivative, evidencing a change in permeability.

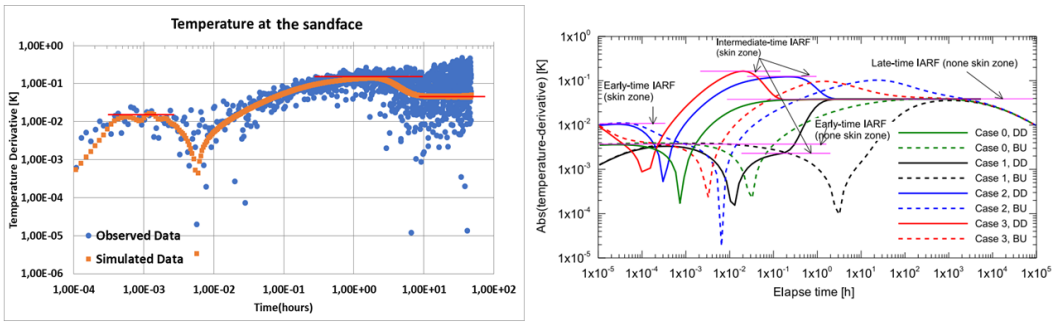

Figure 5.15: The graph on the left is the Bourdet derivative considering the simulated temperature data (in orange) and the observed data (in blue) that is the simulated data in addition to the white noise durign the draw dwon (DD). The graph on the right was presented by Onur et al.[8], in which has the temperature derivative for differents scenarios with some demarcated permeabilities levels.

The following boxplots contain a comparison of the estimation made for Case-1 considering only pressure data (pressure) and combined data considering $(Cdp/Cdt)(1e-4)$ and $(ID/EJTO)(1e-4)$ as $\mathbf{C}_D$ matrix. For the single pressure analysis was assumed a standard deviation of 10Bar[4], and there was no normalization of observed data or $\mathbf{C}_D$ matrix.

Figure 5.16 on the left contains the boxplot for skin permeability (K_1), and on the right, the boxplot for permeability outside the skin zone (K_2). The red line marks the expected values in the plots.

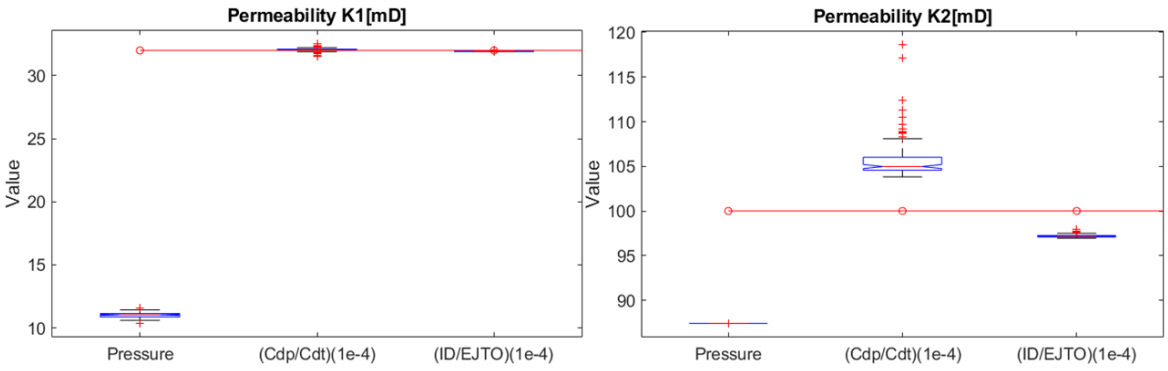

Figure 5.16: Comparison of estimates made considering only pressure data with those made using combined data. On the left we have the estimated of the skin permeability, and on the right we have the permeability outside of the skin region.

The result for the estimation of the permeability of the skin region considering only the pressure data was discrepant from the expected (identified by the red line with markers). Due to analysis made on the pressure plot at the beginning of this chapter, this poor estimative was already expected. In the graphical analysis, the pressure data did not show clear evidence of the identification of this change in permeability.

However, when combined pressure and temperature data is used, the permeability estimation for the skin region is closer to the expected value. For the second permeability, even estimating a value far from ideal for the first permeability, the analysis considering only pressure data estimated an error close to 10%. For the combined data, the error was around 5% in both configurations of the $\mathbf{C}_D$ matrix.

Figure 5.17 presents, on the left side, the boxplot for the α_{skin} estimates, which refers to the radius of the skin region. On the right side, we have the boxplot for the porosity of the reservoir.

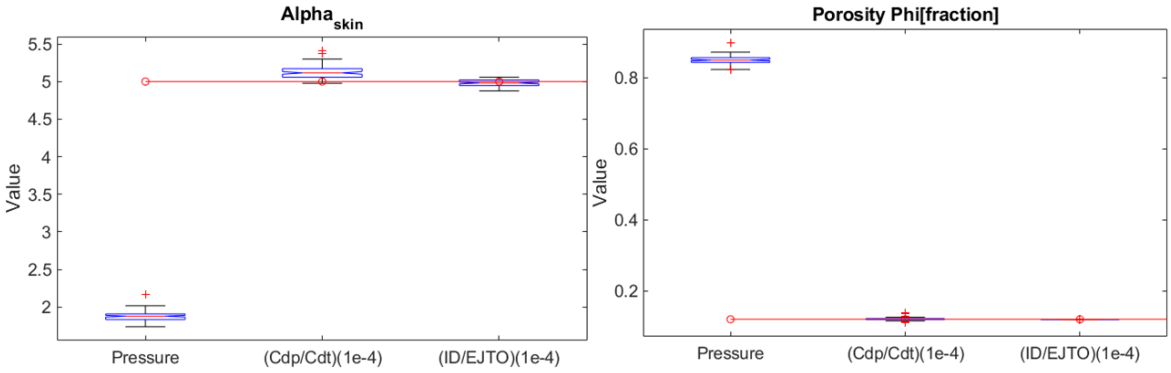

Figure 5.17: Comparison of estimates made considering only pressure data with those made using combined data. On the left we have the estimated of the alpha skin, and on the right we have the porosity of the reservoir.

For both parameters shown in Fig.5.17, the analysis considered only the pressure data gives values far from expected. When combined data is considered, better estimations are obtained for both parameters and in both configurations of the $\mathbf{C}_D$ matrix.

Figure 5.18 illustrates the results of the simulated time series for the final ensemble of parameters originated considering only the pressure data as observed data. The time series originated when the initial ensemble of parameters inserted into the flow simulator is in gray, which are spread out. The circle in red is the observed data, and in blue are the time series from the ensemble that originated after four assimilations of the ensemble ES-MDA.

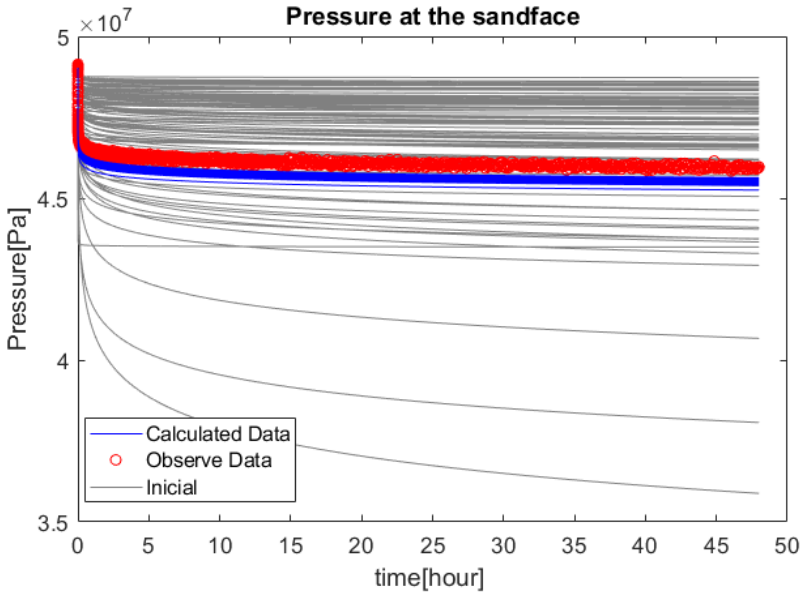

Figure 5.18: Comparison of the pressure evolution originated with the initial ensemble of parameters (in gray) for Case-1. With the observed data (in red) and with the calculated profiles with the final ensemble of parameters (in blue).

It is noteworthy that even not estimating accurately the values for the properties of the skin region and for porosity, the data calculated after the assimilation was close to the observed data. This proximity in the analysis made considering only the pressure data makes the reservoir characterization process and its subsequent management difficult, as the adjustment with the observed data is not representing a good characterization in the region close to the well.

Figure 5.19 contains the time series results for the analysis that considers pressure and temperature as observed data. The configuration of the $\mathbf{C}_D$ matrix adopted for these results was $(ID/EJTO)(1e-4)$. In blue are the results for temperature and pressure after 4 assimilations from the ES-MDA, in red is the observed data and in grey are the results when the initial ensemble is inserted in the flow simulator. We can highlight that for this $\mathbf{C}_D$ matrix configuration, the calculated data pass through the middle of both observed data points, which indicates an accurate data matching, and as the curves in blue are not very spread out, we have a small uncertainty associated with the estimated parameters.

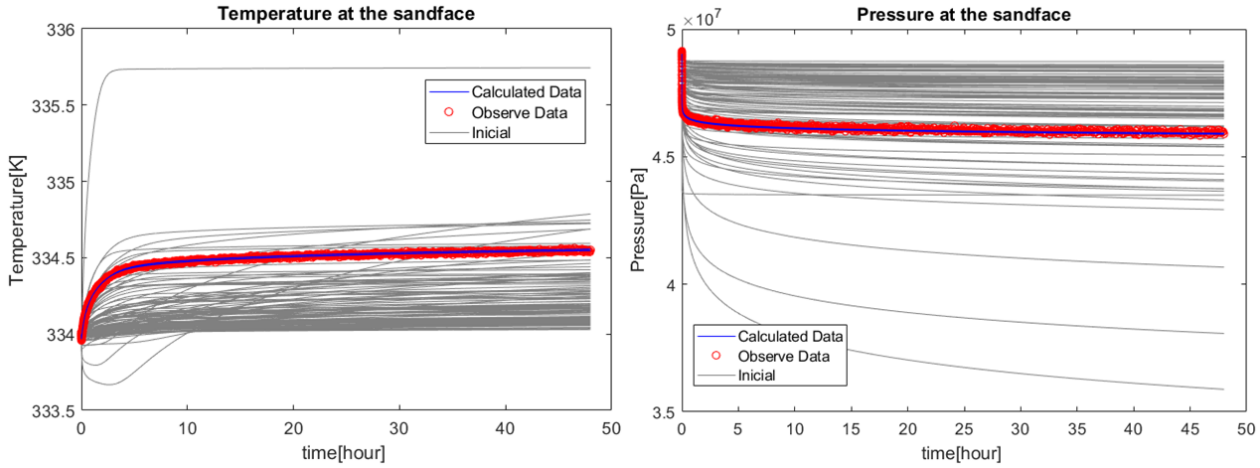

Figure 5.19: Comparison between the pressure and temperature evolution originated with the initial set of parameters (in gray) for Case-1. With the observed data (in red) and with the calculated evolution (in blue) with the final set of parameters. The observed data considered as a couple of pressure and temperature with the matrix $\mathbf{C}_D$ which contains the value of the Joule Thomson coefficient with 1e-4 as weight.

Figure 5.20 contains the analysis's time series results that consider pressure and temperature as observed data. The configuration of the $\mathbf{C}_D$ matrix adopted for these results was $(Cdp/Cdt)(1e-4)$. For this $\mathbf{C}_D$ matrix configuration, the calculated data is located in the middle of the temperature observed data and deviated a little in pressure observed data. Compared to the previously presented configuration ($(ID/EJTO)(1e-4)$), the uncertainty in the parameter estimation is a little bit higher, which leads to an increase in the scattering of the blue curves.

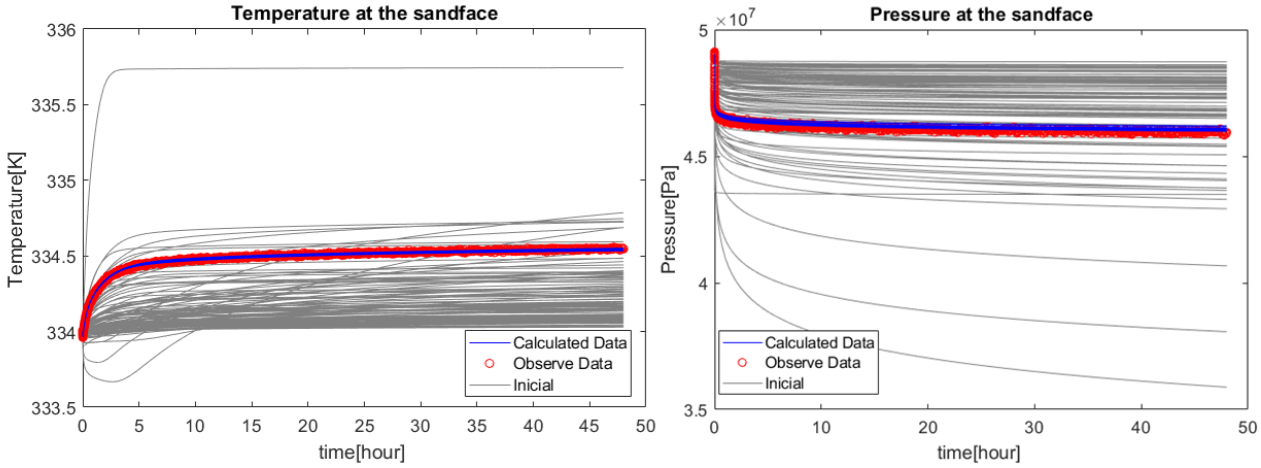

Figure 5.20: Comparison of the pressure and temperature evolutions from the initial set of parameters (in gray) for Case-1. With the observed data (in red) and with the calculated profiles (in blue) with the final set of parameters. The observed data considered as a couple of pressure and temperature with the matrix $\mathbf{C}_D$ which contains the normalized literature values with 1e-4 as weight.

Figure 5.21 summarizes the parameters estimated in each analysis for Case-1. The plots were created using the median of the final set of each parameter. Using this type of representation, it became apparent that the estimate made considering only the pressure data is not able to characterize the skin area and leads to significant errors even for the region away from the well.

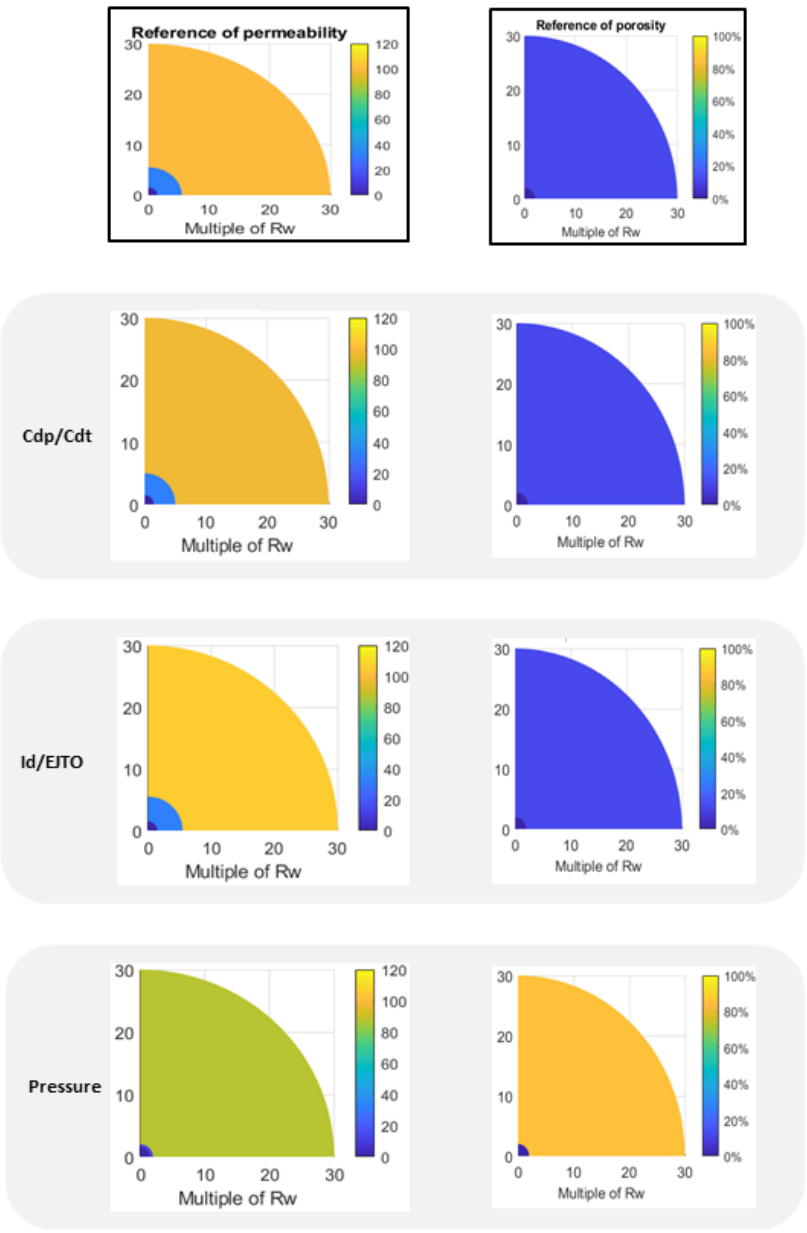

Figure 5.21: Summaries with the medians of the parameters estimated in each analysis for Case-1.

A measure used to evaluate the estimate for each vector $\mathbf{m}$ that makes up the ensemble is the relative error (RE^m) described by the equation below.

$$RE^m = \sum_{i=1}^{Nm} \left[\frac{|Value_{calculated}(i) - Value_{expected}(i)|}{Value_{expected}(i)} \right] \tag{5-5}$$

The index i is used to iterate the parameter vector $\mathbf{m}$, and N_m is the number of parameters inside the vector. The sum of RE^m is calculated and multiplied by 100%, according to equation 5-6, to evaluate the performance of the ensemble.

$$PE = \left[\sum_{j=1}^{N_e} RE^m \right] \times 100\% \tag{5-6}$$

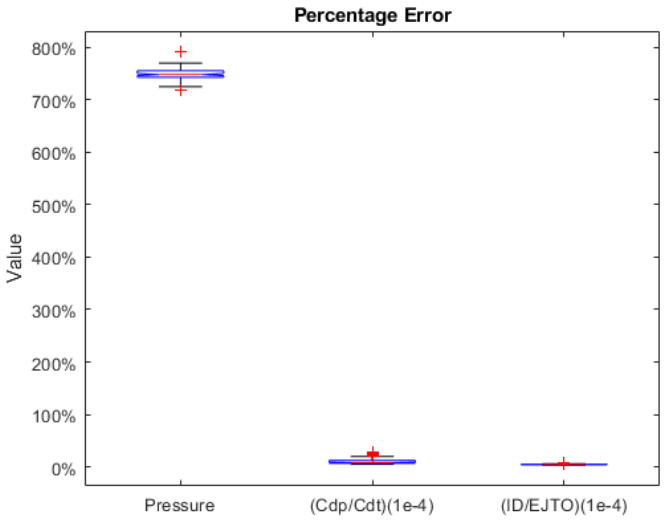

Figure 5.22: Percentage error for the entire ensemble in each analyses for the Case-1.

Figure 5.22 indicates that the percentage error (PE) for the analysis carried out with just the pressure data is about 700 times greater than the PE obtained using combined pressure and temperature data. Suggesting that the addition of the temperature data improved the characterization of the reservoir.

For a complete analysis of Case-1, the root mean square error(RMSE) of the parameters estimated in each analysis are presented, defined by the equation below.

$$RMSE = \left[\sum_{N=1}^{Nm} \frac{|Value_{calculated} - Value_{expected}|^2}{Nm} \right]^{1/2} \tag{5-7}$$

The RMSE was calculated for the permeability of the skin, for the permeability outside the skin region, and the RMSE was also calculated for both permeabilities together, described by the formulation below:

$$\begin{aligned} RMSE_{Permeability-Total} = \Bigg[& \sum_{N=1}^{Nm} \frac{|K1_{calculated} - K1_{expected}|^2}{Nm} \\ & + \sum_{N=1}^{Nm} \frac{|K2_{calculated} - K2_{expected}|^2}{Nm} \Bigg]^{1/2} \end{aligned} \tag{5-8}$$

Finally, was also calculated the RMSE for the reservoir porosity and the alpha skin. All results are presented in the table below.

RMSE	$Pressure$	$(Cdp/Cdt)(1e-4)$	$(ID/EJTO)(1e-4)$
Skin-Permeability	20.982	0.1376	0.0796
Out Skin-Permeability	12.593	6.3432	2.806
Permeability-Total	24.471	6.3447	2.8071
Porosity	3.1279	0.1475	5.0676×10^{-2}
alpha-skin	0.7287	3.5137×10^{-3}	1.1355×10^{-3}

Table 5.1: Table of the RMSE for each parameter in each analysis, for Case-1.

According to table 5.1, the estimation for the skin permeability and the porosity had a big responsibility for the large error in estimating parameters in the single pressure analysis.

5.4 Case-2

Figure 5.1 shows that Case-2 has an intermediate region between the skin region, and the homogeneous region extending to the reservoir's boundary. For this case, we are assuming that we already know the radius of the skin region, so the vector $\mathbf{m}$ is:

$$\mathbf{m} = [log(K_1),\ log(K_2),\ log(K_3),\ log(alpha_{dist})\ and\ \phi]^T \tag{5-9}$$

The values K_1, K_2, and K_3 in vector $\mathbf{m}$ (eq.(5-9)), represent the permeability values in the skin region, intermediate region, and homogeneous region, respectively. The parameter α_{dist} was made considering the same relationship presented in equation 5-1, but for this case, the relationship is between the radius of the well and the radius of the intermediate region.

$$r_{intermediate\ zone} = \alpha_{dist} \times r_w \tag{5-10}$$

The intermediate region has an external radius of 2.5 meters, which leads to a value for variable α_{dist} of 16.03. We also used the $\log(\alpha_{dist})$ as a target.

The porosity (ϕ) of the reservoir continues to be constant in all the reservoir domain. The initial ensemble for the log permeabilities and the $\log(\alpha_{dist})$ was created using a normal distribution with a mean equal to 5.0 and a standard deviation equal to 1.0 ($N(5.0, 1.0^2)$). For the porosity, it was used a normal distribution with a mean of 0.2 and a standard deviation of 0.05 ($N(0.2, 0.5^2)$), the initial ensemble is shown in Fig. 5.23.

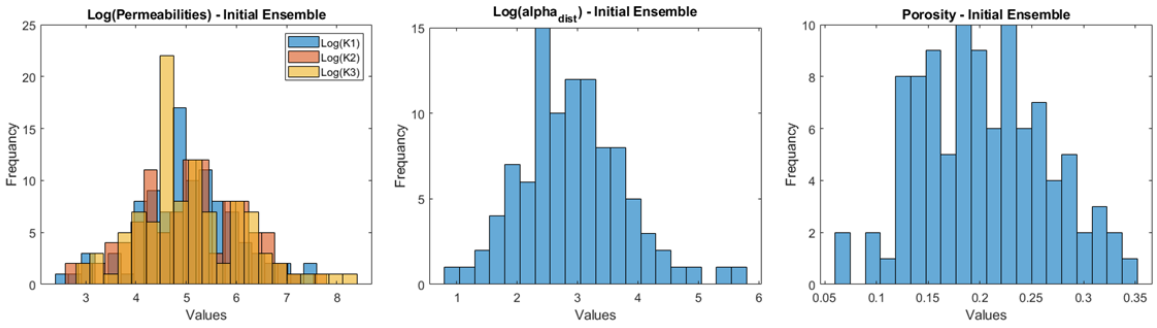

Figure 5.23: Initial distributions for the variabels in analysis into the Case-2.

Case-2 has the data from 4 days of constant flow ($800m^3/day$) to identify all the permeabilities levels in the temperature derivative. Figure 5.24 is the temperature and the pressure time series for the simulated data (orange curve), which are the solutions of the direct problem, and the observed data (blue curve), obtained by adding noise to the simulated data.

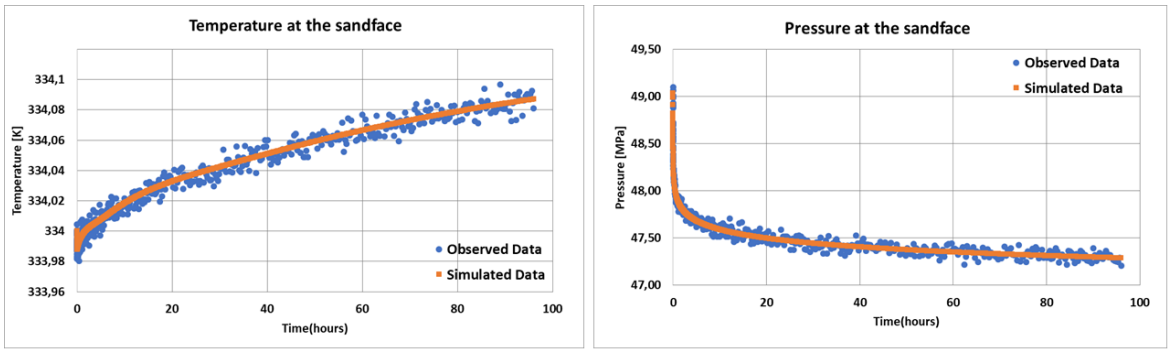

Figure 5.24: Temperature and pressure time series for the Case-2. In orange is simulated data and in blue the observed data.

Figure 5.25 contains the semi-log plots for temperature and pressure data. The pressure semi-log presents a single slope, which indicates a single permeability value. The high permeability can explain this effect near the well that contributes to accelerating the pressure diffusion process. In this way, information from the farthest parts of the well is quickly transmitted. After the intermediate zone, the reservoir assumes a homogeneous characterization of 100mD, the semi-log pressure analysis shows the information from this homogeneous region.

However, analyzing the temperature semi-log of the simulated data and having the notion that this case has 3 distinct regions we can see that it has 3 different slopes, indicated by the red lines. With the temperature semi-log, we

are able to extract the information that the Case-2 reservoir has three different permeabilities. This fact can be explained by the slow velocity of the thermal diffusion, making it possible to identify each permeability present in Case-2.

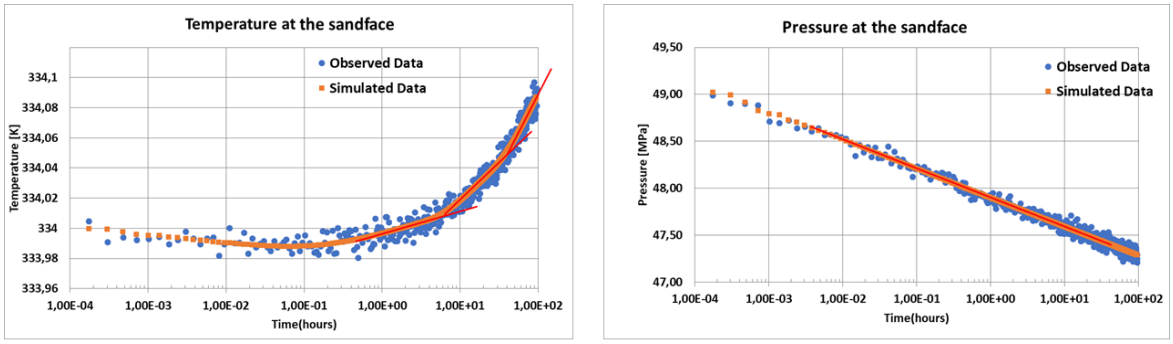

Figure 5.25: Semi-log analysis of the temperature and pressure evolutions for the simulated and observed data, with the slopes demarcated with the red lines.

Figure 5.26 shows the log-log plots for the temperature and pressure derivatives for the simulated and observed data. To approximate the shapes of the curves of the observed data with the simulated data, and facilitate the identification of the levels, the interval of points used in the derivatives calculus was changed. In the pressure log-log, it is evident that when we perform this technique to approximate the data, we start to lose some of the information at an early time, which can make the graphic analysis process even more difficult.

Analyzing the derivatives plots, shows the importance of using temperature data. In the temperature derivative, The three levels of different permeabilities are noticeable in the temperature derivative, while in the pressure derivative, only a single level is presented.

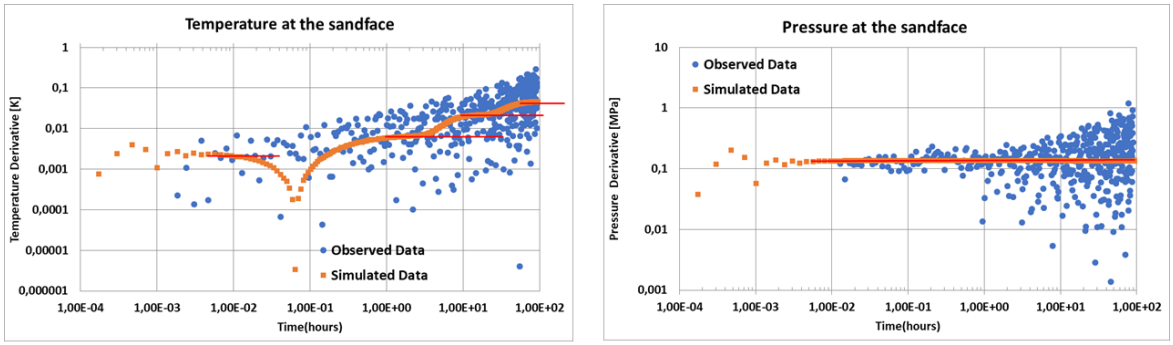

Figure 5.26: Temperature and pressure derivative for the simulated and observed data, with the derivateive levels demarcated.

Figures 5.27 to 5.29, present the boxplots results of each parameter's estimation of Case-2. Those estimations were obtained considering $(ID/EJTO)(1e-4)$ and $(Cdp/Cdt)(1e-4)$ as $\mathbf{C}_D$ matrices for the coupling observed data, and considering only pressure observed data. Figure 5.27 contains the boxplot for permeability K_1 (skin permeability) on the left, and on the right, the boxplot for permeability K_2 (permeability of the intermediate zone).

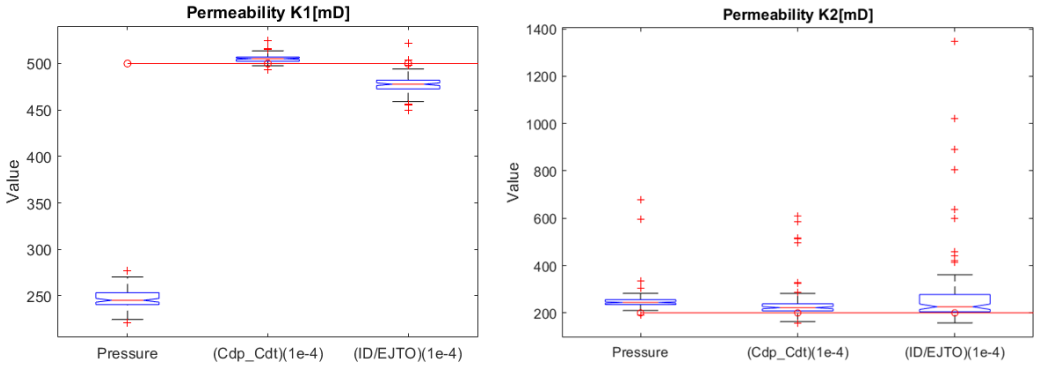

Figure 5.27: Comparison of estimates made considering only pressure data with those made using combined data. On the left we have the estimated of the skin permeability, and on the right we have the permeability of intermediate zone.

Similarly to Case-1, the solution of the inverse problem using only pressure data cannot make an accurate estimation for the permeability of the skin region. When the analysis is made using the combined data, the accuracy of the skin permeability estimation is improved. For the permeability of the intermediate zone (K_2), using only pressure data lead to estimations values close to the values estimated by the combined data. However, the coupled data's medians are still closest to the expected value.

Figure 5.28 contains the boxplots for the estimation of the homogeneous permeability (K_3) on the left and the α_{dist} on the right. The pressure data can identify the homogeneous region. Pressure analysis on the left side of Fig.5.28. The single pressure analysis made an accurate estimation for the permeabilty of the homogeneous region. The analyses using combined data also made accurate estimations with less than 5% of error between the median of the boxplot and the expected value.

Despite making an accurate estimation for the homogenous permeability, the single pressure analysis did not accurately estimate the transition radius, compared to the analyses using combined data. α_{dist} expected value was 16.03, and the median of the estimations considering coupled data was almost 15

while the median of the estimation considering only pressure data was a bit bigger than 20.

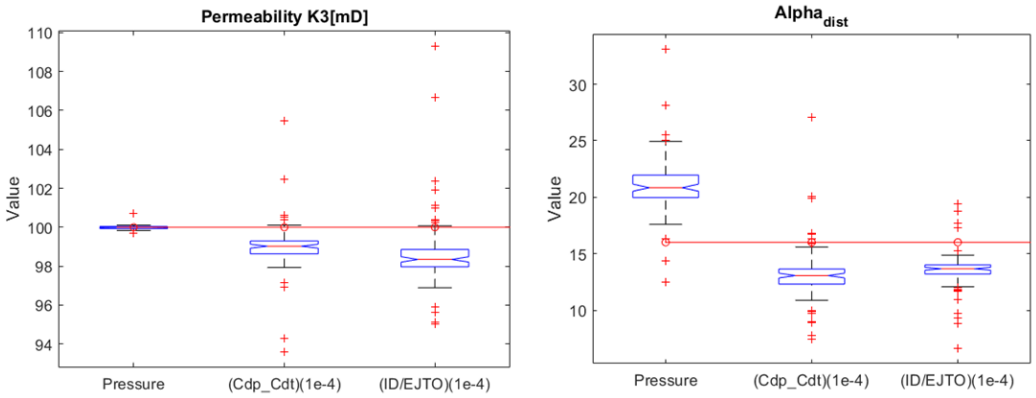

Figure 5.28: Comparison of estimates made considering only pressure data with those made using combined data. On the left we have the estimated of the third permeability, and on the right we have the estimated of the α_{dist}.

Figure 5.29 contains the boxplots for the estimation of the porosity. Similar to the results shown for Case-1, the analysis made considering only the pressure data cannot accurately estimate the reservoir's porosity. On the other hand, the analyses made with the combined data lead to an accurate estimation of the porosity.

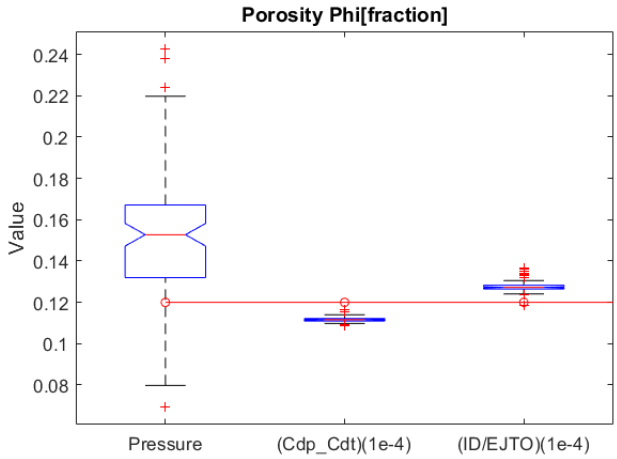

Figure 5.29: Comparison of porosity estimations made considering only pressure data with those made using combined data.

Figure 5.30 contains the percentage error of each analysis made for Case-2. In this case, the percentage error of the single pressure analysis is higher than the coupled analyses, and according to the increase of the complexity of

the reservoir configuration, the distribution of the percentage error from each ensemble also increases. The increase in the error percentage is also associated with the higher parameter estimation uncertainty from the vector **m**.

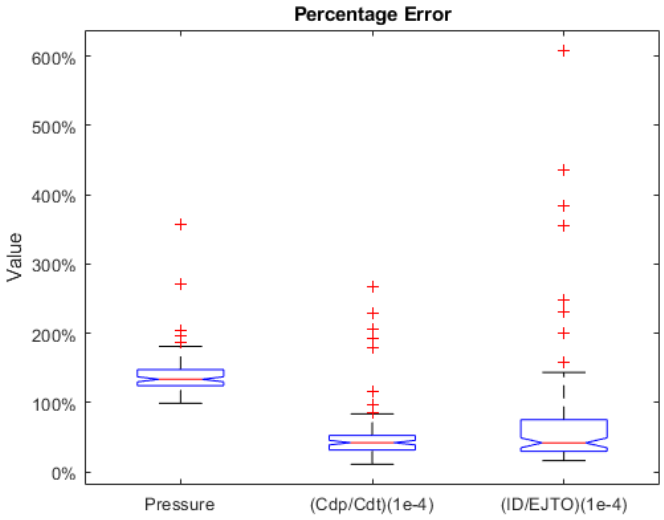

Figure 5.30: Percentage error for the entire ensemble in each analyses for the Case-2.

Figures 5.31 to 5.33 contain the results of the data matching between the observed data (red circle) and the pressure and temperature data originated with the final ensemble. Figure 5.31 shows the results made into the analysis that considers only pressure as observed data. The calculated data (blue lines) passes by the observed data's center. This behavior is explained by the precise adjustment of the permeability of the homogeneous region since the skin and intermediate regions do not significantly affect the behavior of the pressure transient curve. To compare with the initial distribution of parameters, Fig.5.31 also shows in gray the pressure curves originated with the initial ensemble of parameters.

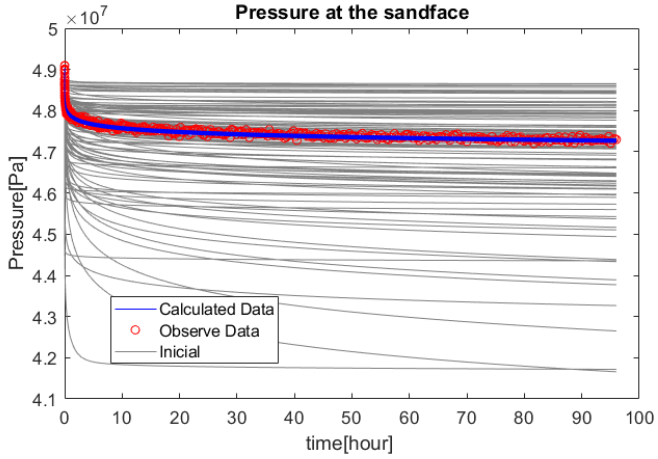

Figure 5.31: Comparison of the pressure evolution originated with the initial set of parameters (in gray) for Case-2. With the observed data (in red) and with the calculated profiles with the final set of parameters.

Figure 5.32 compares the observed data and the calculated temperature and pressure transient data. This analysis was considering the final ensemble originated using $(ID/EJTO)(1e-4)$ as $\mathbf{C}_D$ matrix. Observing the scattering from the initial curves, grey lines, and the final scattering (blue curves) the ES-MDA produces a successful regularization. The scattering from the blue curves is associated with the uncertainty in the parameter estimation.

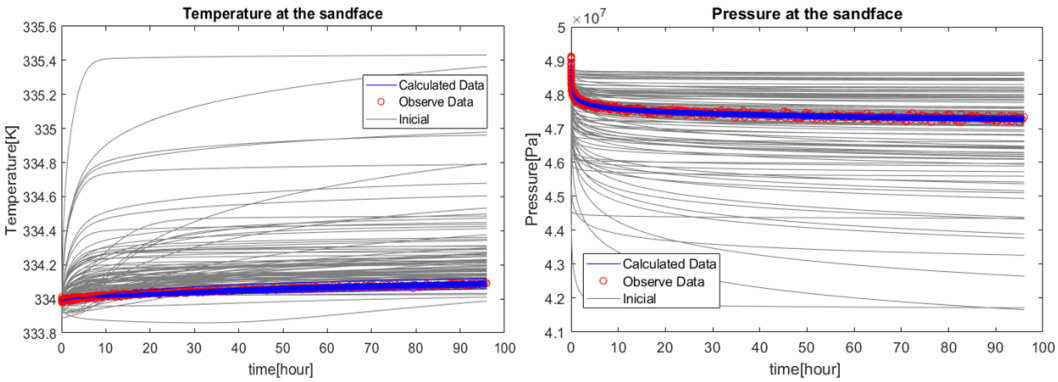

Figure 5.32: Comparison of the pressure and temperature evolution originated with the initial set of parameters (in gray) for Case-2. With the observed data (in red) and with the calculated evolution (in blue) with the final set of parameters. The $\mathbf{C}_D$ configuration used to generate this result was $(ID/EJTO)(1e-4)$.

Figure 5.33 shows the comparison between the observed data and the calculated temperature and pressure transient data. This analysis was considering the final ensemble originated using $(Cdp/Cdt)(1e-4)$ as $\mathbf{C}_D$ matrix.

We can observe that this $\mathbf{C}_D$ matrix configuration also provides a satisfactory regularization.

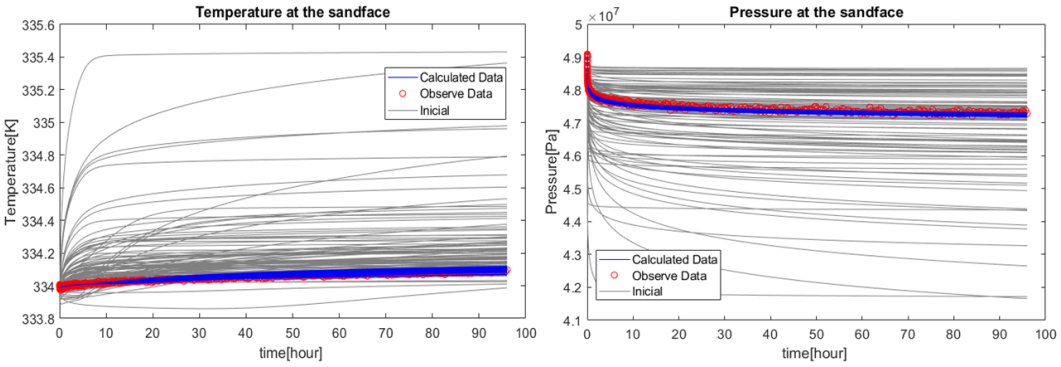

Figure 5.33: Comparison of the pressure and temperature evolutions from the initial set of parameters (in gray) for Case-2. With the observed data (in red) and with the calculated profiles (in blue) with the final set of parameters. The $\mathbf{C}_D$ configuration used to generate this result was $(Cdp/Cdt)(1e-4)$.

In this case, the matching with the observed data from the analysis made considering only pressure data was accurate. However, the skin region was not well characterized even with its pre-defined radius using oly pressure data. Figure 5.34 contains the summary with the mean value of the estimates for each analysis. In this summary, it is evident that the skin's permeability is not well characterized using oly pressure data, resulting in a model with only 2 regions and not 3 regions. But using coupled data, we was able to characterize the three reriongs of the resevoir.

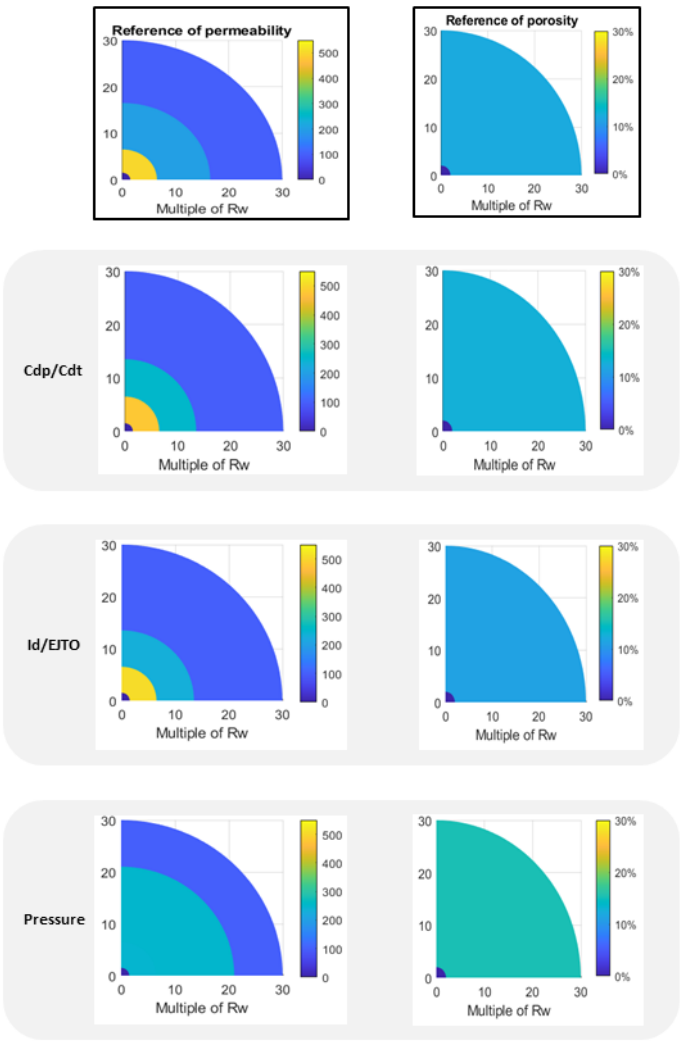

Figure 5.34: Summaries with the medians of the parameters estimated in each analysis for Case-2.

Table 5.2 contains the RMSE of the parameters analyzed in Case-2. This table shows, one more time, that the pressure estimates are less accurate for the permeability of the skin region and porosity. For the other parameters, the accuracy with the pressure data is similar to the combined data.

RMSE	$Pressure$	$(Cdp/Cdt)(1e-4)$	$(ID/EJTO)(1e-4)$
Skin-Permeability	253.82	6.608	24.108
Transition-Permeability	80.171	86.407	192.21
Homogeneous-Permeability	0.101	1.5549	2.2335
Permeability-Total	266.18	86.673	193.73
Porosity	0.0467	8.392×10^{-3}	8.0271×10^{-3}
alpha-Distance	5.556	3.6479	2.9113

Table 5.2: Table of the RMSE for each parameter in each analysis, for Case-2.

5.5 Case-3

As presented in Fig. 5.1, Case-3 also has an intermediate zone after the skin region, but in this case, the intermediate region is more significant than in case-2, with an external radius of 100 meters.

For this case, the parameter vector $\mathbf{m}$ has 6 variables. The first three parameters are the log permeabilities, the following two are related to the heterogenous permeability radius and the skin radius, and the last one is the porosity of the reservoir.

$$\mathbf{m} = [log(K_1),\ log(K_2),\ log(K_3),\ log(r_K),\ \alpha_{skin}\ and\ \phi]^T \tag{5-11}$$

For this analysis, the variable α_{skin} continues to be considered with a relationship between the well's radius and the skin region's radius, as shown in eq.(5-1). The r_K is the external radius from the intermediate zone.

Figure 5.35 shows the distribution of variables for the initial set used in this analysis. For porosity, we considered a normal distribution with a mean 0.2 and a standard deviation of 0.05 ($N(0.2, 0.05^2)$). The other parameters were considered a normal distribution with a mean of 5 and a standard deviation of 1 ($N(5.0, 1.0^2)$).

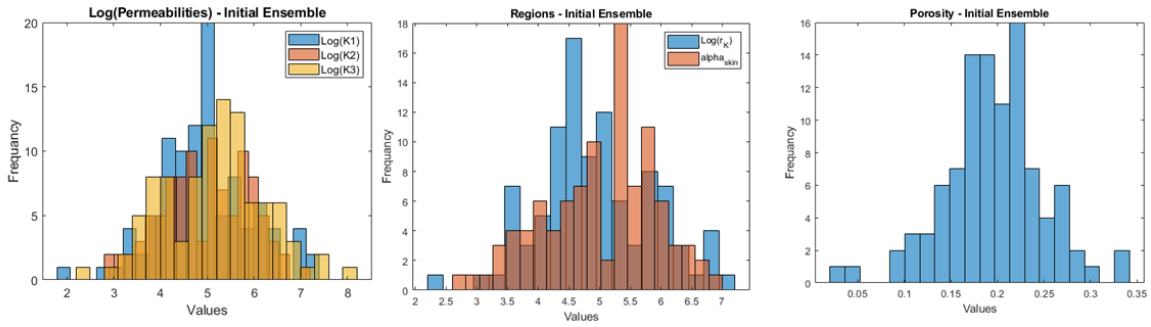

Figure 5.35: Initial distributions for the variabels in analysis into the Case-3.

Figure 5.36 shows the temperature and the pressure times series for the simulated data, which results from the direct problem (orange), and the observed data (blue), considering 4 days of a constant flow rate of $800 m^3/day$.

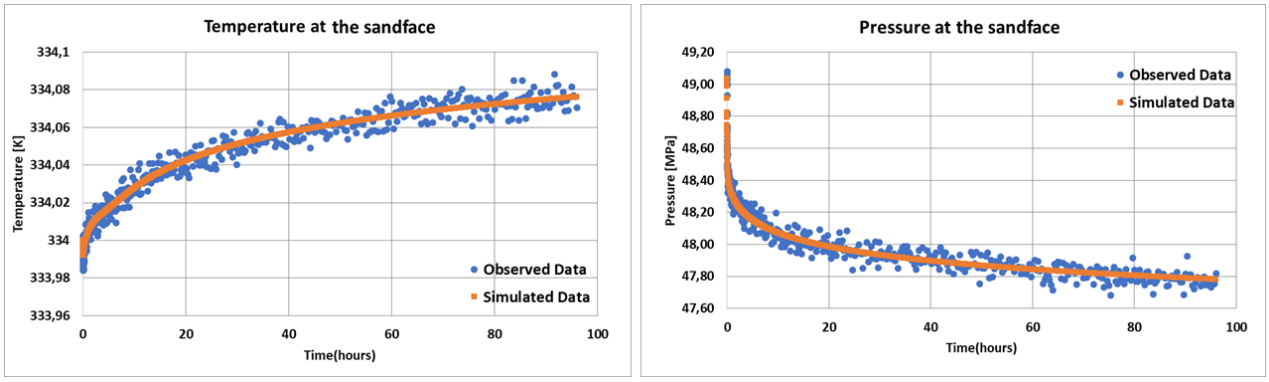

Figure 5.36: Temperature and pressure evolution for the Case-3. In orange is simulated data and in blue the observed data.

Figure 5.37 shows the semi-log plots for temperature and pressure data. Analyzing the pressure semi-log, we observe two slopes representing different permeabilities values. Analyzing the temperature semi-log we can see that it also has 2 different slopes.

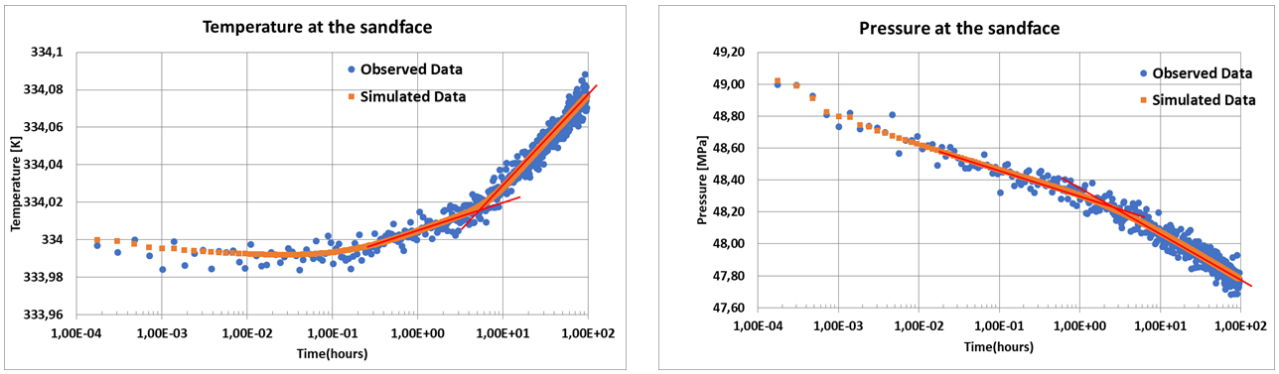

Figure 5.37: Semi-log analysis of the temperature and pressure evolutions for simulated and observed data, with the slopes demarcated with the redlines.

As the reservoir configuration that was analyzed in case-3 has 3 different permeability values, it means that there is a region in which both data can indicate the value through the semi-log plot. According to the results of the previous cases, the temperature is indicating the permeability of the skin region and the intermediate region. The pressure indicates the permeabilities furthest from the well, the intermediate, and the final homogeneous region. So for this reservoir configuration, both data provide information about the intermediate region.

Figure 5.38 contains the temperature and pressure derivative plots where two levels are identified in each plot.

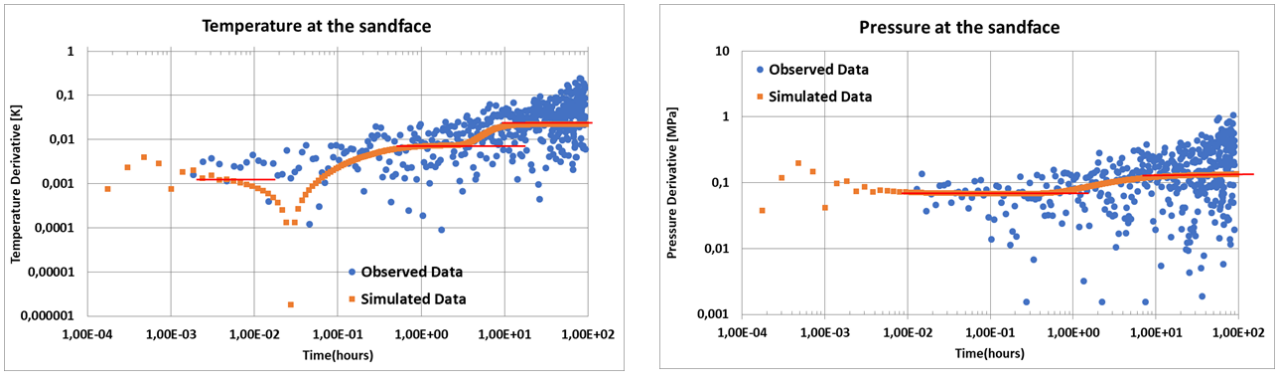

Figure 5.38: Temperature and pressure derivative for the simulated and observed data, with the derivateive levels demarcated.

Figures 5.39 to 5.41 presents the boxplots of the parameter estimation, comparing estimations made considering only the pressure data and the estimations made using the coupled pressure and temperature data considering $(Cdp/Cdt)(1e-4)$ and $(ID/EJTO)(1e-4)$ as $\mathbf{C}_D$ matrix.

Figure 5.39 contains the estimates for the skin permeability (K_1) and the α_{skin}, which is the variable correlated with the skin radius. The figure shows that the estimation made considering only the pressure data is far from the expected value in both parameters. However, when we use the combined pressure and temperature data to estimate, we have an accurate characterization of the skin region.

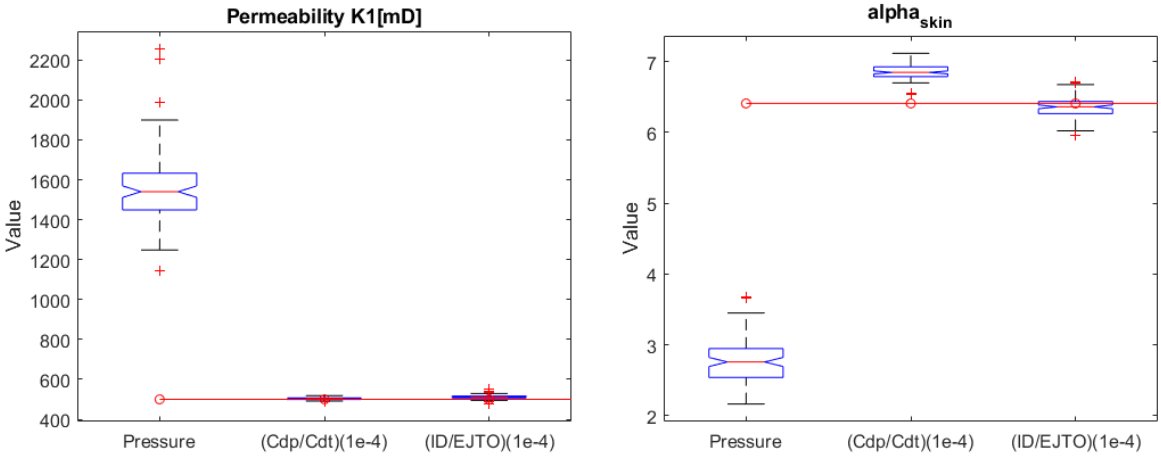

Figure 5.39: Comparison of estimates made considering only pressure data with those made using combined data. On the left we have the estimated of the skin permeability(K_1), and on the right we have the α_{skin}.

Figure 5.40 contains the characterization of the intermediate region of the reservoir, which is composed of the K_2 permeability and the radius r_K. For K_2 permeability, the estimations using $(ID/EJTO)(1e-4)$ achieve the values closest to the expected value, but the estimations using only pressure data, which got farther from the target, only missed by approximately 15mD. For the r_K estimation, the single pressure analysis produces the best estimations. Among the estimates made using coupled data, the one made considering $(ID/EJTO)(1e-4)$ produced the best estimates.

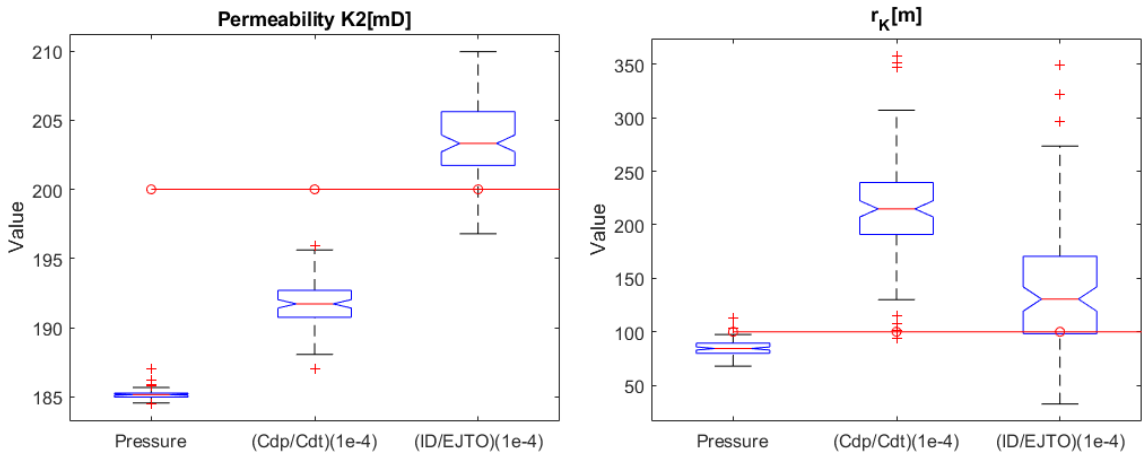

Figure 5.40: Comparison of estimates made considering only pressure data with those made using combined data. On the left we have the estimated of the permeability from the intermediate zone, and on the right we have the external radius of the intermediate zone.

Figure 5.41 contains the estimations of the homogeneous permeability K_3 and the porosity. Considering only the pressure data for analysis, K_3 permeability estimate was more accurate than when using combined pressure and temperature data. The lower accuracy when considering coupled data can be explained since this third permeability does not impact the temperature data that are used in the estimation process. However, as the pressure data is coupled, we can still obtain an accurate estimation when compared to the initial distribution of this permeability. We see that the combined data still obtain estimations more accurately than when considering only the pressure data for the porosity estimation.

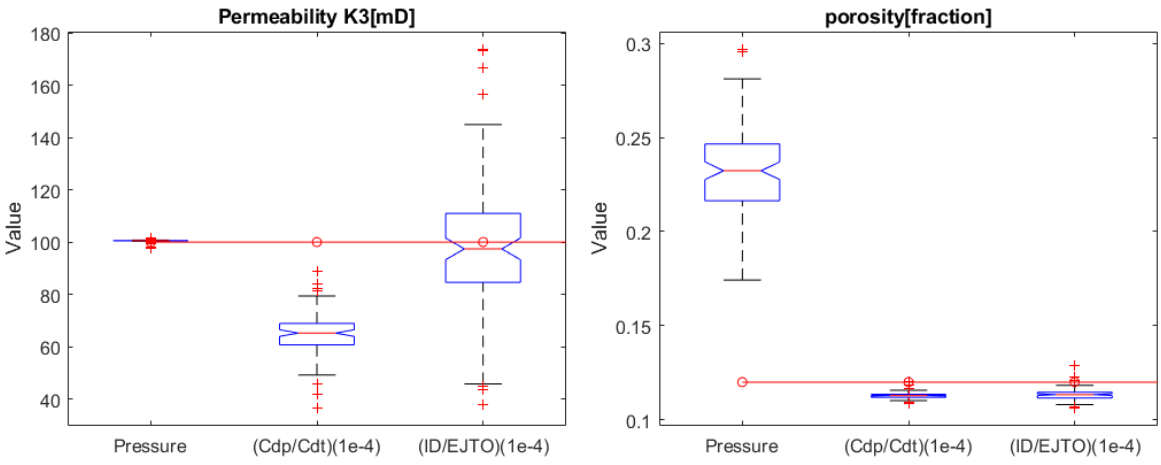

Figure 5.41: Comparison of estimates made considering only pressure data with those made using combined data. On the left we have the estimated of the permeability from the homogeneous zone, and on the right we have the porosity of the reservoir.

Figure 5.42 contains the percentage error of each analysis made for Case-3. We see that although we have some variables in which the estimation using only pressure data is more accurate than those produced with coupled data, the percentage error, considering all the paramters of vector **m**, for the single pressure analysis still higher than the coupled analyses.

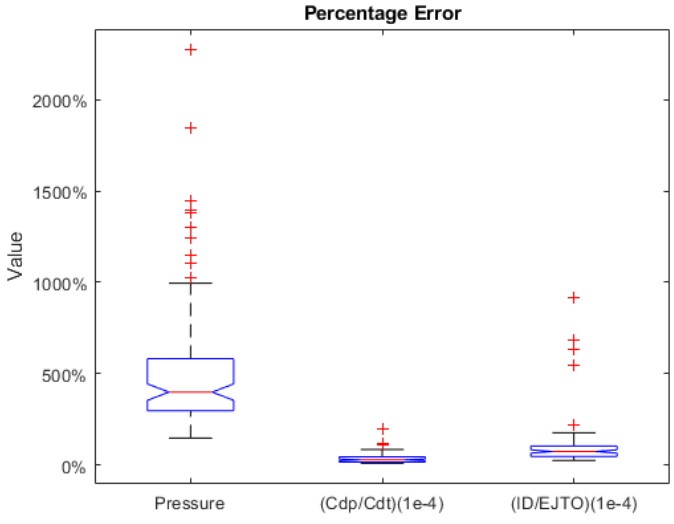

Figure 5.42: Percentage error for the entire ensemble in each analyses for the Case-3.

Figure 5.43 illustrate the results of the simulated time series for the final ensemble of parameters originated considering only the pressure data

as observed data. The time series originated when the initial ensemble of parameters inserted into the flow simulator is in gray, which are spread out. The circles in red are the observed data, and in blue are the time series from the ensemble that originated after four assimilations of the ensemble ES-MDA.

Analyzing Fig.5.43 we can see the ES-MDA method produces accurate data matching with the observed data even without getting a good characterization for the skin region.

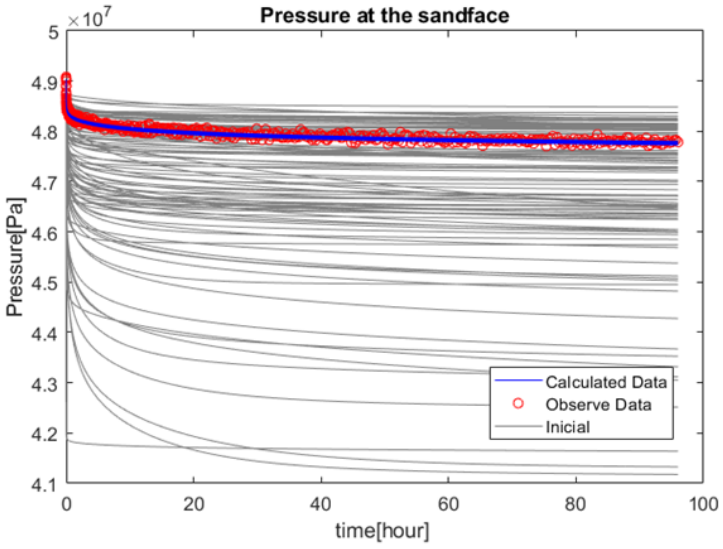

Figure 5.43: Comparison between the pressure evolution originated with the initial set of parameters (in gray) for Case-3. With the observed data (in red) andwith the calculated profiles with the final set of parameters.

Figure 5.44 illustrates the results of the simulated time series for the final ensemble of parameters originated considering coupled data with $(ID/EJTO)(1e-4)$ as $\mathbf{C}_D$ matrix. Due to the homogeneous permeability estimation has a wide range of values, the calculated pressure time series (blue curves) has a wide spread when compared to the analysis of previous cases.

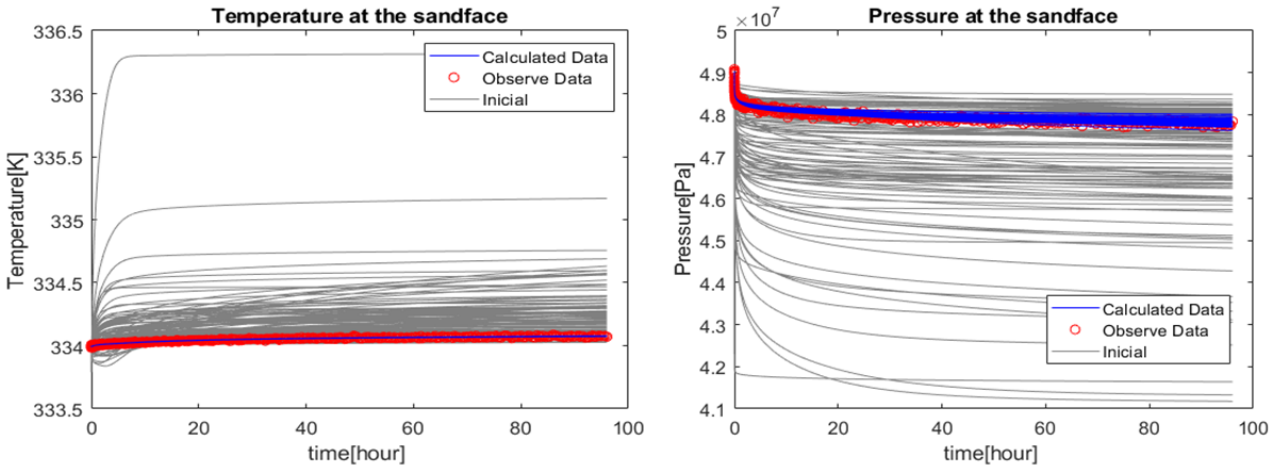

Figure 5.44: Comparison of the pressure and temperature evolution originated with the initial set of parameters (in gray) for Case-3: observed data (in red), the calculated evolution (in blue), and the final ensemble of parameters. The $\mathbf{C}_D$ configuration used to generate this result was $(ID/EJTO)(1e-4)$.

Figure 5.45 illustrates the results of the simulated time series for the final ensemble of parameters originated considering coupled data with $(Cdp/Cdt)(1e-4)$ as $\mathbf{C}_D$ matrix. In this analysis, we can also see the scattering of the calculated pressure curve. However, when compared to the results in Fig.5.44 we can observe that the scattering of the pressure curves is slightly smaller. It can be justified because the range of estimated values for the homogeneous permeability (Fig.5.41) is smaller using coupled data for this $\mathbf{C}_D$ matrix configuration.

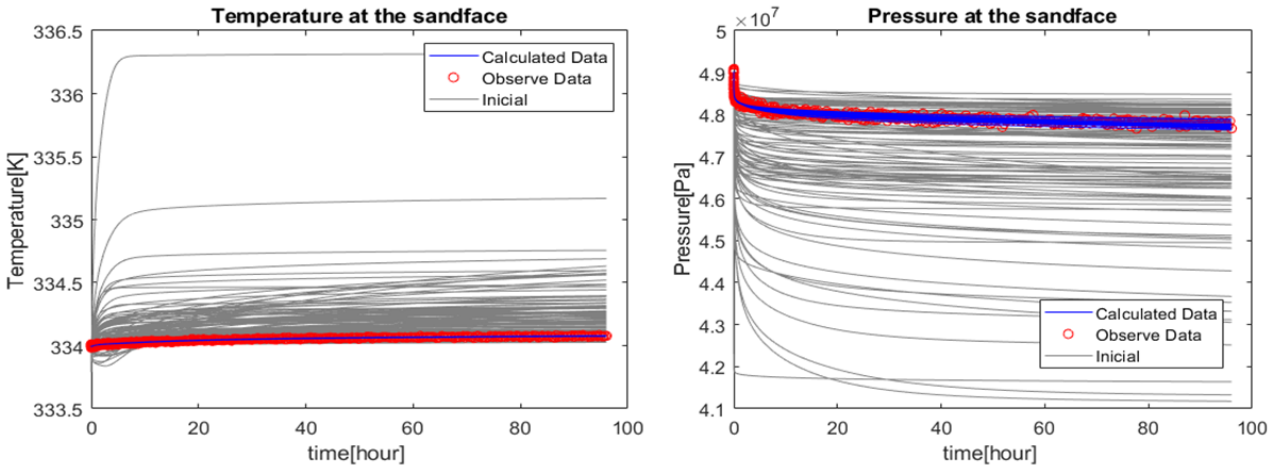

Figure 5.45: Comparison of the pressure and temperature evolutions from the initial set of parameters (in gray) for Case-3: observed data (in red), the calculated profiles (in blue), and the final set of parameters. The C_D configuration used to generate this result was $(Cdp/Cdt)(1e-4)$.

Figure 5.46 summarizes the mean value of the estimation for each analysis. Because the intermediate zone, in this case, is much larger than the skin region, the summary below needed to split the permeability characterization into two columns. The first column is the characterization of the skin region and the permeability of the intermediate zone. The second column is the external radius and the permeability of the intermediate zone, as well as, the permeability of the final homogeneous region. The third column is the porosity of the reservoir.

Figure 5.46 shows clearly that the analysis considering only the pressure data cannot correctly characterize the region near the well (first column), unlike the analyzes made considering coupled data. For the characterization of the transition and homogeneous region, the analysis using only pressure data produces the closest estimation, followed by the analysis made with $(ID/EJTO)(1e-4)$. Finally, considering the porosity estimation, the single pressure data estimates values far from the expected.

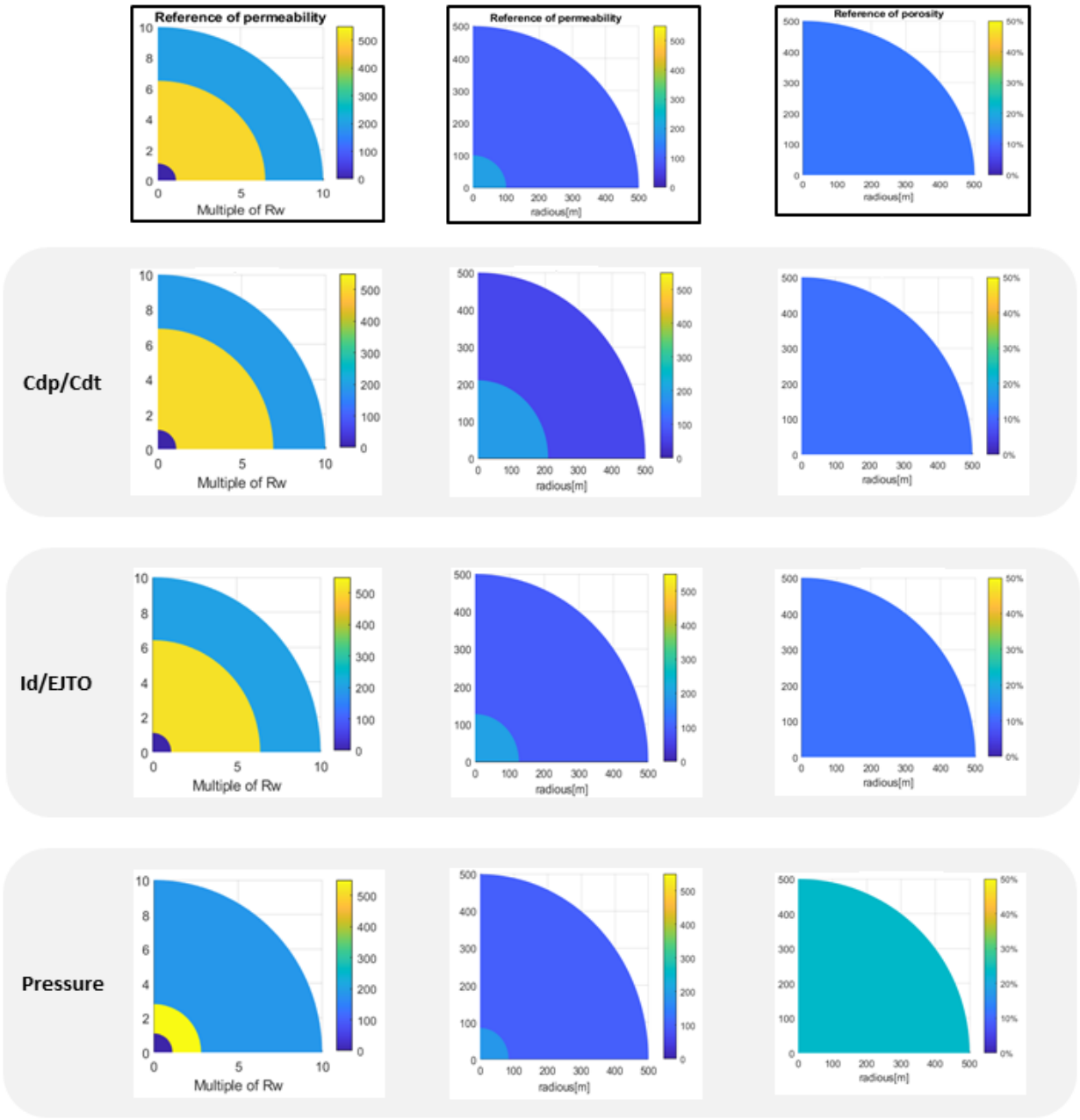

Figure 5.46: Summaries with the medians of the parameters estimated in each analysis for Case-3.

Table 5.3 contains the RMSE of parameters' estimation for Case-3. In this table, the estimations considering only pressure data better estimate r_K and homogenous permeability. On the other hand, this same analysis obtains the worst estimations for the skin region, porosity, and intermediate permeability.

RMSE	Pressure	(Cdp/Cdt)(1e-4)	(ID/EJTO)(1e-4)
Skin-Permeability	1070.7	6.3165	15.898
Transition-Permeability	14.833	8.4304	4.3883
Homogeneous-Permeability	0.652	36.191	25.676
Permeability-Total	1070.8	37.693	30.517
Porosity	0.114	7.324×10^{-3}	7.364×10^{-3}
α_{skin}	3.644	0.458	0.142
r_K	16.863	124.7	74.397

Table 5.3: Table of the RMSE for each parameter in each analysis, for Case-3.

5.6 Case-1-Tidal Effects

The presence of sinusoidal pressure oscillation can be observed during some well test. The amplitude of those oscillations is around 1 psi and their frequency suggests the tidal effect [[73],[74]]. Although these periodic oscillations do not affect the oil production process, they have an important influence on pressure data interpretation and can lead to errors during static tests and pulse tests. It is well known that during a pressure buildup period test, the additional pressure change due to tidal effects can be large enough so that it is difficult to determine information about flow regimes from the late-time buildup data [75].

In order to improve the information content of the computed pressure derivative, the tidal component is usually removed from the measured pressure data. Levitan and Phan [75] used signals from either tidal potential function or seafloor pressure to separate the tidal signals from the pressure data. However, most of the times, not all the tidal components are filtered completely, which leaves a few harmonic components in the pressure data. Due to unfiltered tidal signals, there are misinterpretations in well test analysis [76]. Therefore, there is a challenge in estimating reservoir properties without filtering these periodic disturbances using inverse problem methods.

In this section, the pressure and temperature transient data are modified by adding of a white Gaussian noise in order to represent noise coming from the sensor equipment. To simulate the tidal effect, it was also added a harmonic noise. After that, the inverse problem was solved using the ES-MDA method to characterize the reservoir using the artificially modified pressure and temperature data, taking into account different sources of noise coming from the equipment and the tidal effect (pressure data only). Reservoir configuration

represented by Case-1 is used in the analysis.

$$d_{dobs} = Simulated\ Data + White\ Noise + Harmonic\ Noise \tag{5-12}$$

Different magnitudes of the harmonic noise were considered. A base amplitude of $50KPa$ was defined, and later multiples of this value were tested, as shown in table 5.4.

	White	Tidal
Clean Data (Red circle)	0.0 KPa	0.0 KPa
White+2xTidal (Black circle)	50 KPa	100 KPa
White+4xTidal (Pink circle)	50 KPa	200 KPa
White+6xTidal (Green circle)	50 KPa	300 KPa

Table 5.4: Table with the amplitude of the noise used in each analysis.

Figure 5.47 shows a comparison between a tidal effect of approximately 1 psi amplitude according to Hemala and Balnaves[74] the 50 KPa noise used as a reference for this analysis.

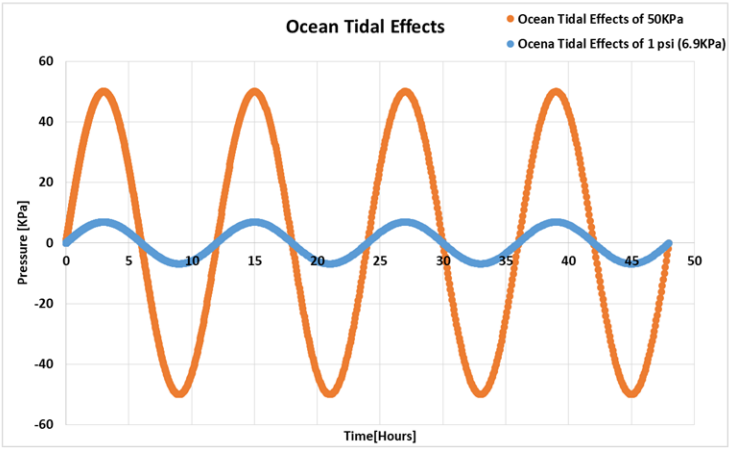

Figure 5.47: Comparison between the harmonic noise with $50KPa$ of amplitude and the harmonic noise with 1 psi.

As illustrated in Fig. 5.47, we are using as reference (orange points) to create the harmonic noise, a value that is about seven times the magnitude found in the literature (blue points) in order to assess whether the ES-MDA would still manage to produce accurate estimates with the combined data.

Figure 5.48 shows the comparisons between the data from the simulator ("clean") with the observed data for the different analyses. As mentioned in the beginning of this section, we can observe that the only changes in the temperature data come from a Gaussian noise with the same amplitude.

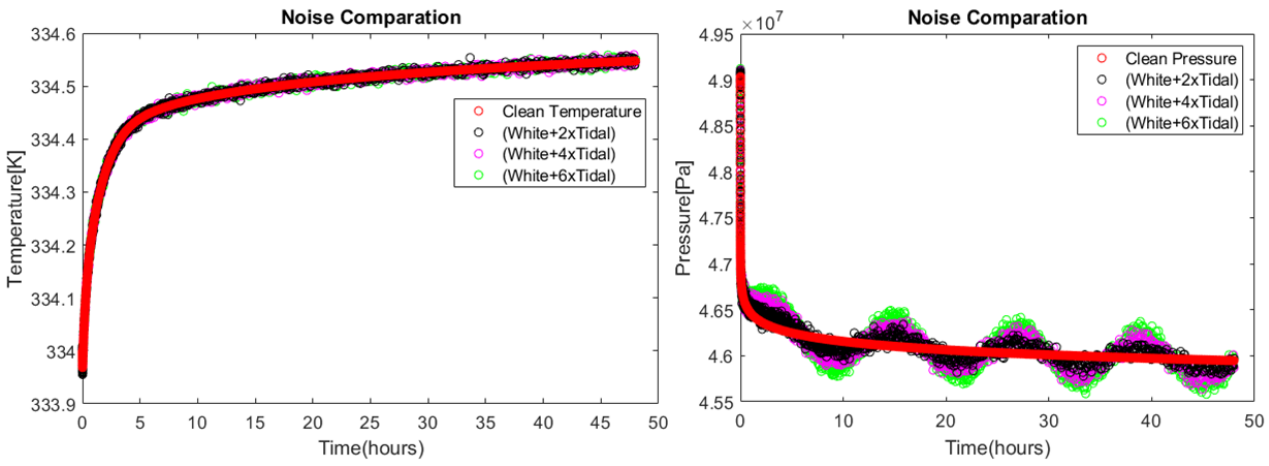

Figure 5.48: Comparison between simulated and observed data with different magnitudes of harmonic noise.

Figure 5.49 shows the graphical analysis for the pressure observed data, in which a harmonic noise of $300KPa$ amplitude was added. In the semi-log analysis of Fig. 5.49, the harmonic noise causes oscillations in the plot, making it difficult to identify its inclination, which can cause errors in this graphical analysis. In the analysis of the pressure derivative, harmonic noise makes it very difficult to identify the level of the derivative, thus needing to increase the intervals used to calculate the Bourdet derivative. As illustrated in the plot on the right side of Fig. 5.49, this increase in interval leads to loss of information over a period of time, making graphical analysis even more difficult.

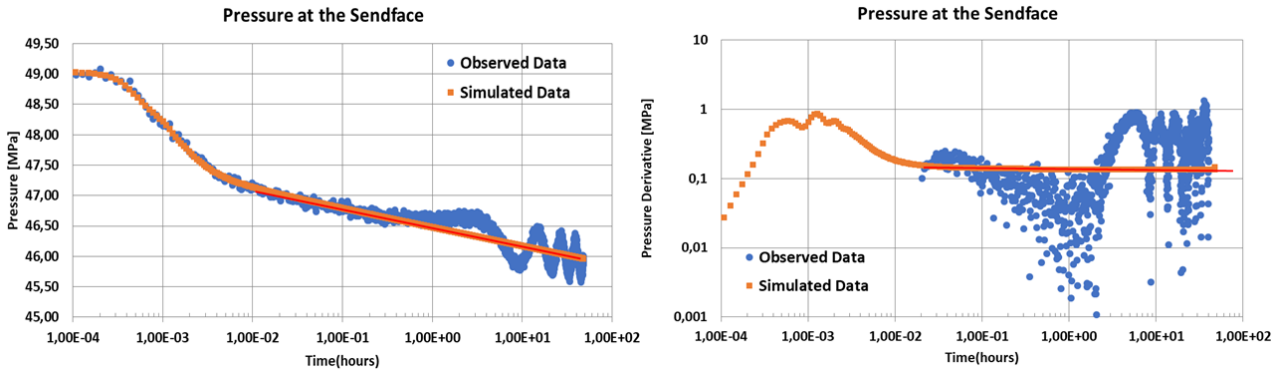

Figure 5.49: On the left is the graph of the semi-log analysis of simulated and observed data for the 'White+6xTidal' analysis. On the right is the pressure derivative graph for that same analysis.

The following plots show the parameter estimations for Case-1 (Fig. 5.1). Estimates are compared when considering only the pressure data ('Pres-

sure(white+0xTidal)') with the combined pressure and temperature data for different multiples of harmonic noise.

For this analysis, the adopted $\mathbf{C}_D$ matrix for the combined data was $(ID/EJTO)(10^{-4})$, and the magnitude of the tidal noise was set: 2, 4, and 6 times the reference value (Fig.5.47), as shown in Fig.5.48.

Starting with the estimations of the permeabilities, Fig. 5.50 presents the estimation for the skin permeability (K_1), on the left, and the permeability outside the skin (K_2), on the right.

For K_1, the addition of noise in the pressure data little affects the estimate, this corroborates the analyzes made earlier in this thesis in which we state that the temperature data are responsible for the characterization of this region closest to the well.

For K_2, the addition of noise in the pressure data significantly affects the permeability estimation of this region. For both estimates shown in Fig. 5.50, the estimations made with the combined data led to more accurate results than when considering only the pressure data, even in the case of higher harmonic noise.

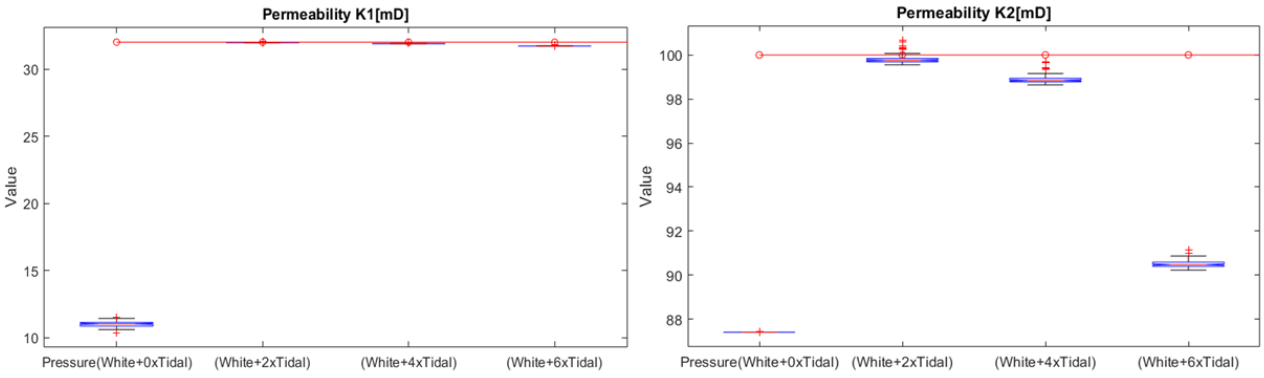

Figure 5.50: Comparison of estimates made considering only pressure data ('Pressure(white+0xTidal)') with those made using combined data with differents amplitudes of harmonic noise.On the left, we have the estimation of the skin permeability, and on the right, we have the permeability out of the skin zone.

Figure 5.51 shows the estimates of porosity. On the right side of Fig. 5.51 is a zoom for the estimates considering the combined data and as perceived for the estimation of the skin permeability. Again, the combined data estimate is closest to the true value than when made considering only the pressure data. The noise in the pressure data had little effect on the porosity estimates.

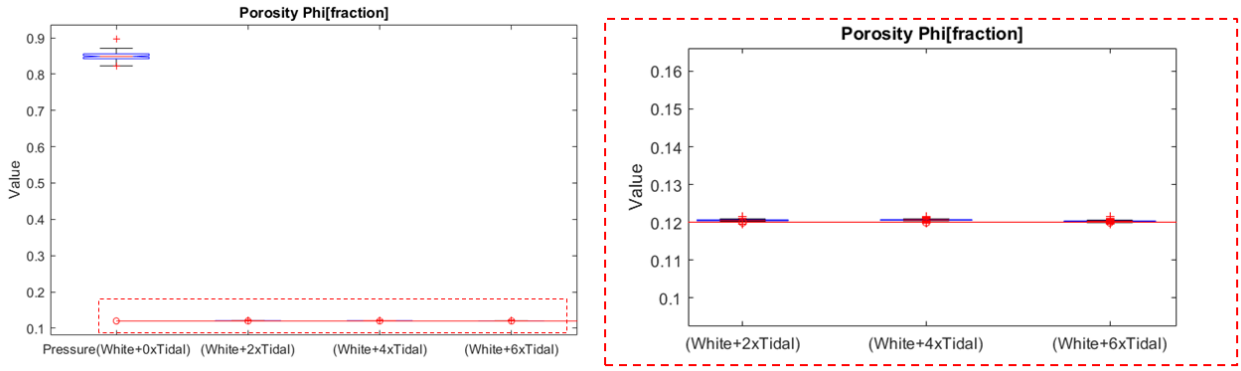

Figure 5.51: Comparison of porosity estimation made considering only pressure data ('Pressure(white+0xTidal)') with those made using combined data with differents amplitudes of harmonic noise. On the right is a zoom in the estimations made by the combined data.

Figure 5.52 presents the estimate of the parameter α_{skin}, and as expected, the addition of noise in the pressure data slightly affected the estimation of α_{skin}.

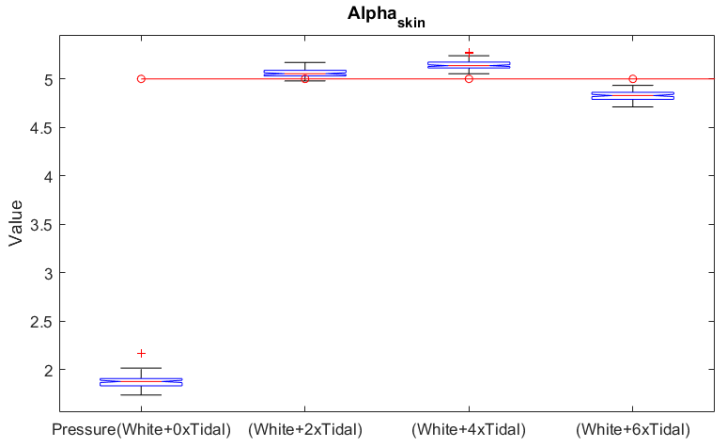

Figure 5.52: Comparison of α_{skin} estimation made considering only pressure data ('Pressure(white+0xTidal)') with those made using combined data with differents amplitudes of harmonic noise.

Figure 5.53 contains the percentage error of each analysis made for Case-1 with the addition of tidal effects in the pressure observed data for the coupled analysis. We see that even with the addition of the tidal effects on the pressure data, the coupling of temperature data in the observed data improves the accuracy of the parameter estimations.

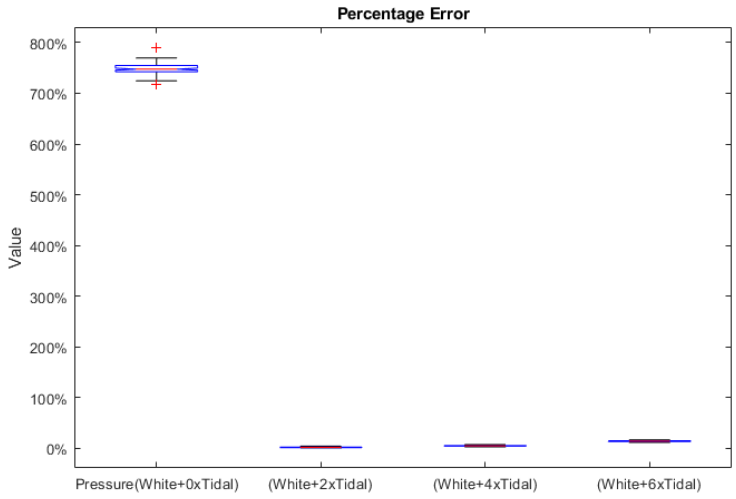

Figure 5.53: Percentage error for the entire ensemble in each analyses for the Case-1 with the addition of tidal effects into the pressure data from the coupled observed data.

The results of the estimates presented in this chapter reinforce that using coupled pressure and temperature data can generate better parameter estimates than when considering only the pressure data, even for cases where there is a significant influence of noise in the measurements.

Figure 5.54 shows the comparison between the observed data, consisting of the temperature and the pressure with the addition of a harmonic noise with 6 times the reference value, and the calculated temperature and pressure transient data. This analysis considered final ensemble originated using $(ID/EJTO)(1e4)$ as $\mathbf{C}_D$ matrix.

Observing the scattering of the blue curves, we see that the estimations of the parameters have a low uncertainty, and as the blue curves pass approximately through the middle of the observed data, we also have an accurate history matching.

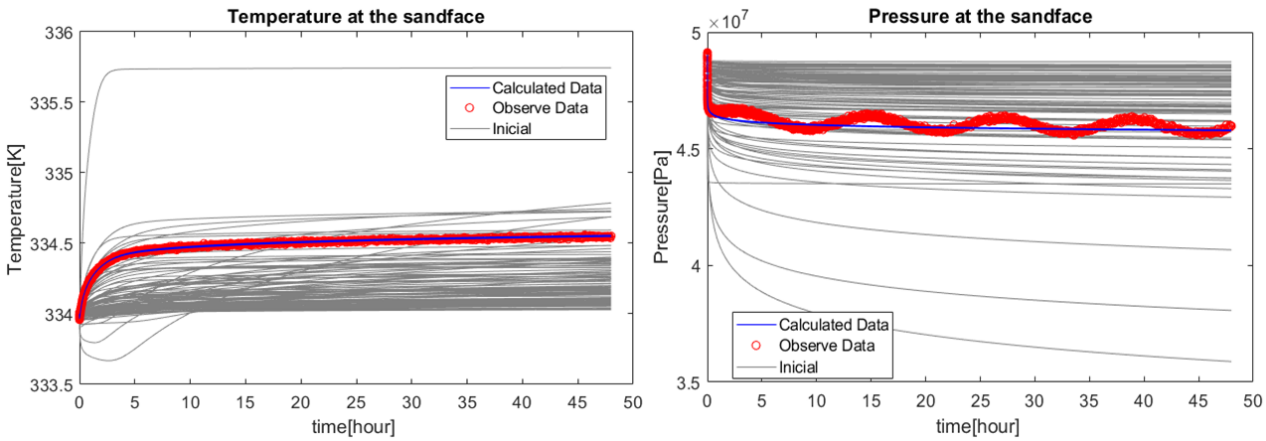

Figure 5.54: Comparison between the pressure and temperature evolutions from the initial ensemble of parameters (in gray) for Case-1 considering tidal effects: observed data (in red), the calculated profiles (in blue), and the final ensemble of parameters. This results were obtaindes considering the coupled data taking and $(ID/EJTO)(1e4)$ as $\mathbf{C}_D$ matrix.

Tables 5.5 and 5.6 contain the RMSE results for each parameter to finalize the analyses for this case. They show that even considering the more significant harmonic noise. The estimate made considering coupled data is still better than the estimate made considering only pressure data without the addition of harmonic noise.

	Pressure (White+0xTidal)	(White+2xTidal)
Skin-Permeability	20.982	0.0328
Out-Skin-Permeability	12.593	0.2769
Permeability-Total	24.471	0.2788
Porosity	3.1279	0.0692
α_{Skin}	0.7287	5.5921e-4

Table 5.5: Table with the RMSE for the parameters considering the tidal effects - part 1

	(White+4xTidal)	(White+6xTidal)
Skin-Permeability	0.1123	0.2843
Out-Skin-Permeability	1.1276	9.4971
Permeability-Total	1.1332	9.5014
Porosity	0.1499	0.1804
$Alpha_{Skin}$	6.627e-4	3.346e-4

Table 5.6: Table with the RMSE for the parameters considering the tidal effects - part 2

6
Conclusions and sugestions

6.1
Conclusions

The main objective of this dissertation was to solve an inverse problem to estimate the parameters of a reservoir and perform an uncertainty analysis of these parameters. The model considered is non-isothermal for a coupled system of a one-dimensional reservoir with a radial flow with the linear well.

The model was validated against commercial software predictions. The model is flexible to make changes in permeability meshes as well as in porosity, but for this work, the porosity was considered constant.

The inverse problem was solved using the ES-MDA method considering different observed data. Normal and harmonic distributions were used to reproduce the data noise from the measuring equipment and the tidal effect. The tests performed for the different observed data showed that it is necessary to be normalized when the observed data are a combination of pressure and temperature data.

Sensibility analysis of $\mathbf{C}_D$ matrix was performed to use the ES-MDA method for combined data. Some values of the components in the $\mathbf{C}_D$ matrix used in this work are found in the literature for uncoupled data analysis.

Evaluating the initial and final parameters distributions presented in the sensitivity analysis of matrix $\mathbf{C}_D$ for the coupled observed data, we can conclude that excluding the configuration of matrix $\mathbf{C}_D$ equal to the identity matrix, the rest of the configurations (Cdp/Cdt, $(Cdp/Cdt)(1e-2)$, $(Cdp/Cdt)(1e{-}4)$, $ID/EJTO$, $(ID/EJTO)(1e{-}2)$, and $(ID/EJTO)(1e{-}4)$) used produced satisfactory estimates.

When comparing the estimates made considering only pressure data with those made with coupled data with $(ID/EJTO)(1e{-}4)$ and $(Cdp/Cdt)(1e{-}4)$ as $\mathbf{C}_D$ matrix, the results show that the estimations with the coupled data are more accurate than the estimation with single pressure data.

Finally, when the harmonic effects (Tidal) are considered in the pressure data, results show that working with coupled data improves the reservoir's characterization and reduces the total percentage error calculated with the

parameter estimates.

6.2 future works suggestions

As suggestions for future work, we can point out new areas for the reservoir-well model and solution of the inverse problem. Regarding the well-reservoir model:

- Implement the model in python;
- Update the model to consider multiphase flow;
- Update the coupled conditions to not only consider full-open wellbore.

For the inverse problem:

- Increase the scale of the reservoir parametrization;
- Use of real oil field production data as observed data, with more than a month of flow;
- Apply the method to multi-layer reservoirs;
- Due to the increase in the number of parameters as well as the use of data from fields, we suggest using the techniques presented by Duarte e Silva [77] to optimize the choice of the inflation factors of the method.

Bibliography

[1] TV Moore, RJ Schilthuis, and William Hurst. The determination of permeability from field data. *Proc., API Bull*, 211(4), 1933.

[2] Fikri J Kuchuk, Mustafa Onur, and Florian Hollaender. *Pressure transient formation and well testing: convolution, deconvolution and nonlinear estimation*. Elsevier, 2010.

[3] Mustafa Onur and Murat Cinar. Analysis of sandface-temperature-transient data for slightly compressible, single-phase reservoirs. *Spe Journal*, 22(04):1134–1155, 2017.

[4] Vinıcius Luiz Santos Silva. *Ajuste de histórico e otimização da produção de petróleo sob incertezas - Aplicação do gerenciamento de reservatórios em malha fechada*. PhD thesis, Universidade Federal do Rio de Janeiro, 2016.

[5] Mustafa Onur, Murat Çinar, et al. Temperature transient analysis of slightly compressible, single-phase reservoirs. In *SPE Europec featured at 78th EAGE Conference and Exhibition*. Society of Petroleum Engineers, 2016.

[6] E Bath. Introduction to well testing. *Schlumberger Wireline & Testing*, 1998.

[7] John Paul Spivey and W John Lee. *Applied well test interpretation*. Society of Petroleum Engineers Richardson, TX, 2013.

[8] Mustafa Onur, Yildiray Palabiyik, O Inanc Tureyen, and Murat Cinar. Transient temperature behavior and analysis of single-phase liquid-water geothermal reservoirs during drawdown and buildup tests: Part ii. interpretation and analysis methodology with applications. *Journal of Petroleum Science and Engineering*, 146:657–669, 2016.

[9] Alain Gringarten. From straight lines to deconvolution: The evolution of the state of the art in well test analysis. *SPE Reservoir Evaluation & Engineering - SPE RESERV EVAL ENG*, 11:41–62, 02 2008.

[10] Silva Junior MF. Intelligent well transient temperature signal reconstruction. 2012.

[11] Mauricio SC Galvao, Marcio S Carvalho, Abelardo B Barreto Jr, et al. A coupled transient wellbore/reservoir-temperature analytical model. *SPE Journal*, 24(05):2–335, 2019.

[12] Mauricio SC Galvao, Marcio S Carvalho, and Abelardo B Barreto Jr. Thermal impacts on pressure transient tests using a coupled wellbore/reservoir analytical model. *Journal of Petroleum Science and Engineering*, 191:106992, 2020.

[13] Weibo Sui, Ding Zhu, Alfred Daniel Hill, and Christine A Ehlig-Economides. Determining multilayer formation properties from transient temperature and pressure measurements. In *SPE Annual Technical Conference and Exhibition*. OnePetro, 2008.

[14] Mustafa Onur and Yildiray Palabiyik. Nonlinear parameter estimation based on history matching of temperature measurements for single-phase liquid-water geothermal reservoirs. In *World Geothermal Congress, Melbourne, Australia*, pages 19–25, 2015.

[15] Yilin Mao and Mehdi Zeidouni. Analytical solutions for temperature transient analysis and near wellbore damaged zone characterization. In *SPE Reservoir Characterisation and Simulation Conference and Exhibition*. OnePetro, 2017.

[16] Filippo Panini, Mustafa Onur, et al. Parameter estimation from sandface drawdown temperature transient data in the presence of a skin zone near the wellbore. In *SPE Europec featured at 80th EAGE Conference and Exhibition*. Society of Petroleum Engineers, 2018.

[17] Celia Maria Ferraz Nakano, Antonio Carlos Capeleiro Pinto, Jose Luiz Marcusso, Kazuioshi Minami, et al. Pre-salt santos basin-extended well test and production pilot in the tupi area-the planning phase. In *Offshore technology conference*. Offshore Technology Conference, 2009.

[18] Jose Ronaldo Chaves Melo, Marcelo Bassetto, Christiane De Camargo, Danillo Duran Camiza, Marcio Alan Svaiter, et al. Ss early testing-jabuti extended well test. In *Offshore Technology Conference*. Offshore Technology Conference, 2009.

[19] L Costa, D Rodrigues, R da Silva, I Souza, MA Fernandes, et al. Intelligent completions used during extended well test of exploratory wells in brazil. In *OTC Brasil*. Offshore Technology Conference, 2017.

[20] Flavio Costa Coccoli, Luiz Carlos Milagre Cruz, Thiago Henrique Galindo Vasconcelos, Rogério Pinto Marti, et al. Intelligent completion in extended well

test. In *Offshore Technology Conference*. Offshore Technology Conference, 2019.

[21] Vinicius Mattoso Reis da Silva. *Estimativa Da Permeabilidade De Reservatórios Com Base Em Dados Transientes De Pressão E Temperatura*. PhD thesis, Undergraduate thesis, Dep. Of Mechanical Engineering, PUC-Rio, 2019.

[22] Willer Planas Gonçalves. *Estimativa de Parâmetros de Reservatórios de Petróleo a Partir de Modelo Transiente não Isotérmico*. PhD thesis, MS Thesis, Dep. Of Mechanical Engineering, PUC-Rio . . . , 2020.

[23] Alexandre A. Emerick and Albert C. Reynolds. Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55:3–15, 2013. Ensemble Kalman filter for data assimilation.

[24] Alexandre A Emerick and Albert C Reynolds. Investigation of the sampling performance of ensemble-based methods with a simple reservoir model. *Computational Geosciences*, 17(2):325–350, 2013.

[25] Luciana dos Santos Silva et al. Integração de análise de incertezas e ajuste de histórico= aplicação em um caso complexo. 2011.

[26] Bohan Xu and Fahim Forouzanfar. The information content and integration of distributed-temperature-sensing data for near-wellbore-reservoir characterization. *SPE Reservoir Evaluation & Engineering*, 20(04):0906–0923, 2017.

[27] AF Van Everdingen. The skin effect and its influence on the productive capacity of a well. *Journal of petroleum technology*, 5(06):171–176, 1953.

[28] AF Van Everdingen and William Hurst. The application of the laplace transformation to flow problems in reservoirs. *Journal of Petroleum Technology*, 1(12):305–324, 1949.

[29] Obinna O Duru, Roland N Horne, et al. Modeling reservoir temperature transients and reservoir-parameter estimation constrained to the model. *SPE Reservoir Evaluation & Engineering*, 13(06):873–883, 2010.

[30] M Sidorova, V Shako, V Pimenov, B Theuveny, et al. The value of transient temperature responses in testing operations. In *SPE Middle East Oil & Gas Show and Conference*. Society of Petroleum Engineers, 2015.

[31] Mustafa Onur, Gonca Ulker, Serhat Kocak, and Ihsan M Gok. Interpretation and analysis of transient-sandface-and wellbore-temperature data. *SPE Journal*, 22(04):1156–1177, 2017.

[32] Obinna O Duru, Roland N Horne, et al. Simultaneous interpretation of pressure, temperature, and flow-rate data using bayesian inversion methods. *SPE reservoir evaluation & engineering*, 14(02):225–238, 2011.

[33] Rim Abdullovich Valiullin, Ayrat Ramazanov, Ramil Faizyrovich Sharafutdinov, et al. Temperature logging in russia: development history of theory, technology of measurements and interpretation techniques. In *Kuwait International Petroleum Conference and Exhibition*. Society of Petroleum Engineers, 2009.

[34] Stephen Prensky et al. Temperature measurements in boreholes-an overview of engineering and scientific applications. *The Log Analyst*, 33(03):313–333, 1992.

[35] Maurice Schlumberger, AA Perebinossoff, and Henri-Georges Doll. *Temperature measurements in oil wells*. [éditeur non identifié][ca. 1936](Bungay, 1936.

[36] Rim Abdullovich Valiullin, Ayrat Ramazanov, Vyacheslav Pavlovich Pimenov, Ramil Faizyrovich Sharafutdinov, Alexandr Sadretdinov, et al. Qualitative and quantitative interpretation: the state of the art in temperature logging. In *North Africa Technical Conference and Exhibition*. Society of Petroleum Engineers, 2010.

[37] A Daniel Hill. Production logging: theoretical and interpretive elements. 1990.

[38] HA Lauwerier. The transport of heat in an oil layer caused by the injection of hot fluid. *Applied Scientific Research, Section A*, 5(2-3):145–150, 1955.

[39] Henry J Ramey Jr et al. Wellbore heat transmission. *Journal of petroleum Technology*, 14(04):427–435, 1962.

[40] EB Chekalyuk. Thermodynamics of oil reservoir. *Nedra, Moscow*, 1965.

[41] PG Atkinson, HJ Ramey Jr, et al. Problems of heat transfer in porous media. In *SPE Annual Fall Technical Conference and Exhibition*. Society of Petroleum Engineers, 1977.

[42] SK Garg and JW Pritchett. On pressure-work, viscous dissipation and the energy balance relation for geothermal reservoirs. *Advances in Water Resources*, 1(1):41–47, 1977.

[43] RN Horne, K Shinohara, et al. Wellbore heat loss in production and injection wells. *Journal of Petroleum Technology*, 31(01):116–118, 1979.

[44] A Sh Ramazanov and VM Nagimov. Analytical model for the calculation of temperature distribution in the oil reservoir during unsteady fluid inflow. *Oil and Gas Business Journal*, 1(1), 2007.

[45] Bulent Izgec, Shah Kabir, A Rashid Hasan, et al. Transient fluid and heat flow modeling in coupled wellbore/reservoir systems. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2006.

[46] Weibo Sui, Ding Zhu, Alfred Daniel Hill, Christine Ehlig-Economides, et al. Model for transient temperature and pressure behavior in commingled vertical wells. In *SPE Russian Oil and Gas Technical Conference and Exhibition*. Society of Petroleum Engineers, 2008.

[47] Jeffrey F App. Nonisothermal and productivity behavior of high-pressure reservoirs. *Spe Journal*, 15(01):50–63, 2010.

[48] MARCIO DA SILVEIRA CARVALHO Maurício da Silva Cunha Galvão and ABELARDO BORGES BARRETO JR. *Analytical Models for Thermal Wellbore Effects on Pressure Transient Testing*. PhD thesis, Master thesis, Dep. Of Mechanical Engineering, PUC-Rio, 2018.

[49] Davut Erdem Bircan and M. Onur. Temperature transient tests: Modeling, interpretation, and nonlinear parameter estimation. 2019.

[50] *Thermal Modeling for Characterization of Near Wellbore Zone and Zonal Allocation*, volume All Days of *SPE Russian Petroleum Technology Conference*, 10 2010. SPE-136256-MS.

[51] Marko Vauhkonen, Tanja Tarvainen, and Timo Lähivaara. *Inverse Problems*, pages 207–227. Springer International Publishing, Cham, 2016.

[52] Albert Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005.

[53] Obinna Onyinye Duru, Roland N Horne, et al. Joint inversion of temperature and pressure measurements for estimation of permeability and porosity fields. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2010.

[54] Alexandre Emerick. *History Matching and Uncertainty Characterization: Using Ensemble-based Methods*. LAP LAMBERT Academic Publishing, 2012.

[55] Filippo Panini, Mustafa Onur, and Dario Viberti. An analytical solution and nonlinear regression analysis for sandface temperature transient data in

the presence of a near-wellbore damaged zone. *Transport in Porous Media*, 129(3):779–810, 2019.

[56] Jose Adriano Bento de Souza Cardoso. *Pressure and Temperature transiente response in a coupled stratified wellbore-reservoir model*. PhD thesis, MS Thesis, Dep. Of Mechanical Engineering, PUC-Rio, 2020.

[57] G Ulker. *Modeling and Interpreting Transient Wellbore Temperature Distributions in Wellbores Under Nonisothermal Single-Phase Liquid Flow Conditions in Oil and Geothermal Reservoirs*. PhD thesis, MS Thesis, Dep. Of Pet. and Nat. Gas Eng., ITU Graduate School of Science . . . , 2016.

[58] G I Barenblatt, V M Entov, and V M Ryzhik. *Theory of fluid flows through natural rocks*. 1 1989.

[59] A Rashid Hasan, C Shah Kabir, Dongqing Lin, et al. Analytic wellbore temperature model for transient gas-well testing. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2003.

[60] Cyril Frank Colebrook, T Blench, H Chatley, EH Essex, JR Finniecome, G Lacey, J Williamson, and GG Macdonald. Correspondence. turbulent flow in pipes, with particular reference to the transition region between the smooth and rough pipe laws.(includes plates). *Journal of the Institution of Civil engineers*, 12(8):393–422, 1939.

[61] Geir Evensen. *Data assimilation: the ensemble Kalman filter*. Springer Science & Business Media, 2009.

[62] Emilio Coutinho. *Assimilação de dados de teste de formação e perfilagem de produção utilizando Filtro de Kalman por conjunto*. PhD thesis, UENF, 04 2011.

[63] Dean S Oliver, Albert C Reynolds, and Ning Liu. *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge, 2008.

[64] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960.

[65] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.

[66] Sigurd I Aanonsen, Geir Nævdal, Dean S Oliver, Albert C Reynolds, and Brice Vallès. The ensemble kalman filter in reservoir engineering–a review. *Spe Journal*, 14(03):393–412, 2009.

[67] Peter Jan Van Leeuwen and Geir Evensen. Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review*, 124(12):2898–2913, 1996.

[68] Jan-Arild Skjervheim, Geir Evensen, et al. An ensemble smoother for assisted history matching. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2011.

[69] Alexandre A Emerick. Analysis of the performance of ensemble-based assimilation of production and seismic data. *Journal of Petroleum Science and Engineering*, 139:219–239, 2016.

[70] Marcel Vinokur. On one-dimensional stretching functions for finite-difference calculations. *Journal of Computational Physics*, 50(2):215–234, 1983.

[71] Mauricio SC Galvao. Analytical models for thermal wellbore effects on pressure transient testing. *Master's thesis, PUC–Rio, Rio de Janeiro*, 2018.

[72] Dominique Bourdet, JA Ayoub, YM Pirard, et al. Use of pressure derivative in well test interpretation. *SPE Formation Evaluation*, 4(02):293–302, 1989.

[73] A.K. Khurana. Influence of tidal phenomena on interpretation of pressure build-up and pulse tests. *The APPEA Journal*, 16(1):99–105, 1976. IPTC-16711-MS.

[74] ML Hemala and C Balnaves. Tidal effect in petroleum well testing. In *Offshore South East Asia Show*. OnePetro, 1986.

[75] Michael M. Levitan and Vinh Phan. Identification of Tidal Signal in Well Test Pressure Data. All Days, 10 2003. SPE-84376-MS.

[76] T Gowtham, G. Moghanloo Rouzbeh, Krishna B Vamsi, and P Srikanth. Possible Misinterpretations in Well Test Analysis Due to Unfiltered Tidal Signal. All Days, 05 2016. SPE-180464-MS.

[77] Sinésio Pesco Thiago M. D. Silva and Abelador borges barreto Jr. *Investigando a geração dos fatores de inflação para o ensemble smoother with multiple data assimilation.* PhD thesis, Ph.D. thesis, Math Departmen, PUC-Rio, 2021.

A
Reservoir Energy - Discretization

Using the linearization procedure and expanding the Darcy's velocity, the energy conservations equation of the reservoir 3-9 is represented by the following equation:

$$\frac{\partial T}{\partial t} - \varphi_t^* \frac{\partial p}{\partial t} = C_{pRo} K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \frac{\partial T}{\partial r} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \alpha_t \frac{\partial T}{\partial r} \right) - C_{pRo} K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \varepsilon_{JTo} \frac{\partial p}{\partial r}$$

Expanding the temporal terms we have:

$$\frac{\partial T}{\partial t} - \varphi_t^* \frac{\partial p}{\partial t} = \frac{T_i^{j+1} - T_i^j}{\Delta t} - \varphi_t^* \frac{p_i^{j+1} - p_i^j}{\Delta t} \tag{A-1}$$

The non-linear part of the partial equation:

$$C_{pRo} K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \frac{\partial T}{\partial r} = C_{pRo} K_i \left(\frac{p_{i+1}^j - p_{i-1}^j}{\Delta rP + \Delta rF} \right) \left[\theta \frac{T_{i+1}^{j+1} - T_{i-1}^{j+1}}{\Delta rP + \Delta rF} + (1-\theta) \frac{T_{i+1}^j - T_{i-1}^j}{\Delta rP + \Delta rF} \right] \tag{A-2}$$

The second order term:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \alpha_t \frac{\partial T}{\partial r} \right) = \frac{\theta}{r_i} \left(\frac{r_r \alpha_t \frac{p_{i+1}^{j+1} - p_i^{j+1}}{\Delta rF} - r_l \alpha_t \frac{p_i^{j+1} - p_{i-1}^{j+1}}{\Delta rP}}{\Delta rM} \right) + \frac{(1-\theta)}{r_i} \left(\frac{r_r \alpha_t \frac{p_{i+1}^{j} - p_i^{j}}{\Delta rF} - r_l \alpha_t \frac{p_i^{j} - p_{i-1}^{j}}{\Delta rP}}{\Delta rM} \right) \tag{A-3}$$

The last term:

$$-C_{pRo} K_i \left(\left. \frac{\partial p}{\partial r} \right|_j \right) \varepsilon_{JTo} \frac{\partial p}{\partial r} = -C_{pRo} K_i \varepsilon_{JTo} \left(\frac{p_{i+1}^j - p_{i-1}^j}{\Delta rP + \Delta rF} \right) \left[\theta \frac{p_{i+1}^{j+1} - p_{i-1}^{j+1}}{\Delta rP + \Delta rF} + (1-\theta) \frac{p_{i+1}^j - p_{i-1}^j}{\Delta rP + \Delta rF} \right] \tag{A-4}$$

B
Wellbore Energy - Discretization

Rewriting equation 2-21 and applying the spatial and temporal derivatives of temperature:

$$\begin{aligned}\rho_o A c_{po}(1+C_T)\left(\frac{T_i^{w,j+1}-T_i^{w,j}}{\Delta t}\right) &= \theta\left[\rho_o Q_i^j c_{po} L_R\left[T_{ext,i}-T_i^{w,j+1}\right]\right.\\ &\quad\left.-\rho_o Q_i^j c_{po}\left(\left(\frac{T_{i+1}^{w,j+1}-T_{i-1}^{w,j+1}}{\Delta Z_N+\Delta Z_S}\right)-\varphi(z,t)+\frac{g\sin(\alpha)}{c_{po}}\right)\right]\\ &\quad+(1-\theta)\left[\rho_o Q_i^j c_{po} L_R\left[T_{ext,i}-T_i^{w,j}\right]\right.\\ &\quad\left.-\rho_o Q_i^j c_{po}\left(\left(\frac{T_{i+1}^{w,j}-T_{i-1}^{w,j}}{\Delta Z_N+\Delta Z_S}\right)-\varphi(z,t)+\frac{g\sin(\alpha)}{c_{po}}\right)\right]\end{aligned} \tag{B-1}$$

The term $\varphi(z,t)$ (eq:2-23), which contains the derivatives of pressure and flow, can be expanded as follows:

$$\varphi(z,t)=\varepsilon_{JTo}\left(\frac{p_{i+1}^w-p_{i-1}^w}{\Delta Z_N+\Delta Z_S}\right)-\frac{Q_i^j}{A^2 c_{po}}\left(\frac{Q_{i+1}-Q_{i-1}}{\Delta Z_N+\Delta Z_S}\right) \tag{B-2}$$

Regarding the temporal discretization "j", it is necessary to be careful, since when the term is multiplied by $(1-\theta)$ it is analyzed at the point "j+1". Except for the flow term, which was used to linearize the formulation.

C
Wellbore Momentum - Discretization

For the moment conservation equation (eq.2-28), the relation of the flow analysis at the time instant "j" will also be applied to linearize the formulation. For the quadratic flow term, one factor will be evaluated at instant "j" and the second will be evaluated according to the θ term. Below is the equation with the expanded derivatives.

$$\frac{1}{A}\left(\frac{Q_i^{j+1}-Q_i^j}{\Delta t}\right)+g\sin(\alpha)=\theta\left[-\frac{Q_i^j}{A^2}\left(\frac{Q_{i+1}^{j+1}-Q_{i-1}^{j+1}}{\Delta Z_{top}+\Delta Z_{bottom}}\right)-\frac{1}{\rho_o}\left(\frac{p_{i+1}^{w,j+1}-p_{i-1}^{w,j+1}}{\Delta Z_{top}+\Delta Z_{bottom}}\right)-\frac{f|Q_i^j|Q_i^{j+1}}{2A^2D}\right]+(1-\theta)\left[-\frac{Q_i^j}{A^2}\left(\frac{Q_{i+1}^j-Q_{i-1}^j}{\Delta Z_{top}+\Delta Z_{bottom}}\right)-\frac{1}{\rho_o}\left(\frac{p_{i+1}^{w,j}-p_{i-1}^{w,j}}{\Delta Z_{top}+\Delta Z_{bottom}}\right)-\frac{f|Q_i^j|Q_i^j}{2A^2D}\right] \tag{C-1}$$