Maximising Oil Production Through Data Modelling, Simulation and Optimisation

José Antonio Peñuelas Alvarez

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Department of Automatic Control and Systems Engineering

March 2017

Acknowledgements

I would like to thank my main supervisor Dr Hua-Liang Wei for his valuable guidance and support during my studies.

I feel privileged for joining the Department of Automatic Control and Systems Engineering. I have learned a lot from all my friends and colleagues I have met at the department. You have all contributed to the amazing time I've had in Sheffield during the last years.

I express my gratitude to Mexico's Ministry of Energy for its financial support, without this help, this work would not have been possible. I would like to thank Dr Pablo Ibargüengoytia for his comments, ideas and monitoring the progress of my studies.

Finally, I would like to thank my family for giving me their love and support. Without you, I would not be here. I dedicate this Thesis to you: Papá, Mamá y Ale.

Abstract

The research work presented on this thesis provides an alternative tool for characterising oil fields under fluid injection by analysing historical production/injection rates. In particular polynomial and radial basis Non Linear Autoregressive with Exogenous Input Model (NARX) models were developed; these models were capable of capturing the dynamics of an operating field in the North Sea.

A Greedy Randomised Adaptive Search Procedure (GRASP) heuristic optimisation method was applied for estimating a future injection strategy. This approach is combined with a risk analysis methodology, a popular approach in financial mathematics. As a result, it is possible to estimate how likely it is to reach a production goal.

According to the simulations, it is possible to increase oil production by 10% in one year by implementing a smart injection strategy with low statistical uncertainty. Resulting from this research project, a computational tool was developed. It is now possible to estimate NARX models from any field under fluid injection as well as finding the best future injection scenario.

Contents

List of Figures

List of Tables

Acronyms

ARX Autoregressive With Extra Inputs Model.

CRM Capacitive Resistive Model.

CRM-P Capacitive Resistive Model Producer-Producer.

EOR Enhanced Oil Recovery.

FROLS Forward Orthogonal Least Squares Algorithm.

GRASP Greedy Randomised Adaptive Search Procedure.

LM Levenberg-Marquardt.

MAPE Mean Absolute Percentage Error.

MIMO Multiple-Input Multiple-Output.

MISO Multiple-Input Single-Output.

MPO Model Predictive Output.

MSE Mean Square Error.

MSRBF Multi-Scale Radial Basis Function.

NARMAX Non Linear Autoregressive Moving Average with Exogenous Input Model.

NARX Non Linear Autoregressive with Exogenous Input Model.

OSA One Step Ahead.

SERR Sum of Error Reduction Ratios.

SISO Single-Input Single-Output.

SRA Statistical Reservoir Analysis.

WEC World Energy Council.

Chapter 1

Introduction

1.1 Background

During the last decades, the demand for energy has been continuously rising, this has been driven by population growth and changes in our modern way of living. According to studies published by BP's statistical review of world energy and the World Energy Council (WEC), fossil fuels are still the main source of energy accounting for almost 90% of the global demand [1], [2]. The following graph shows the distribution of the main energy sources.

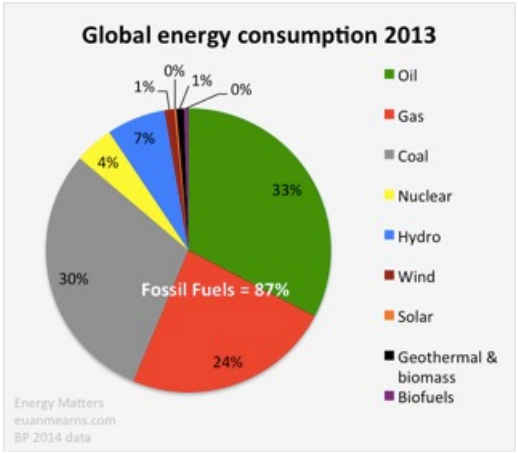

Figure 1.1: Energy Sources

[1]

It is clear that oil and gas supplies are still the dominant sources. Accounting for approximately 57% of the global demand. There is a trend for the development of renewable sources but the requirement of oil and gas fuels is also in growth and it is not likely to change in the near future [2]. It is expected that oil and gas will still be the main sources of energy for most of the current century. The previous figures show the importance of the oil and gas sector worldwide; in some countries, its importance is even more substantial.

The following diagram shows the estimated world energy consumption forecast by energy source.

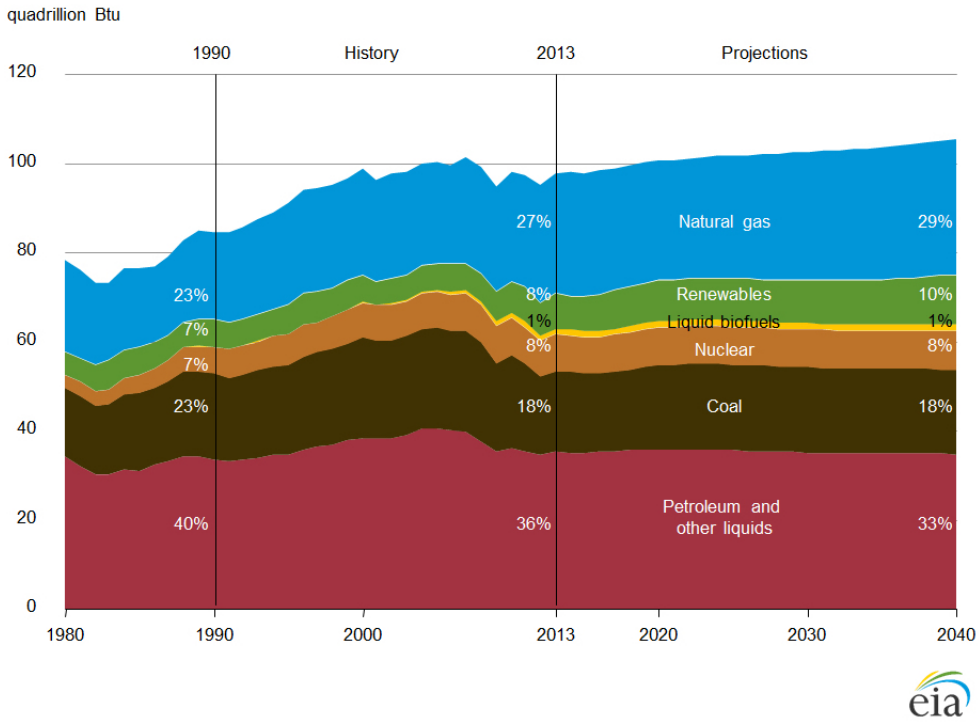

Figure 1.2: World Energy Consumption Forecast

[3]

Oil production can be divided into three different stages. The first one is known as primary recovery. On this stage oil and gas flow into the surface naturally by the reservoir's own initial pressure. Once production from the primary recovery stage hits its peak, normally water, air, CO_2 or nitrogen are injected into the field in order to maintain an adequate pressure level and allow oil to keep flowing into the surface. Once this happens the field is known as a "mature" field, while the artificial pressure lift is known as secondary recovery. Inevitably after some time even with an artificial pressure lift, production will decrease. In some fields heat or chemicals are injected in order to change the oil properties, reducing its viscosity allowing it to

flow more easily. This strategy is known as tertiary recovery [4]. During primary recovery about 10% of the total field reserves can be extracted, for the secondary recovery up to an extra 30% while the tertiary recovery adds up to an additional 30%. Normally the tertiary recovery is not executed since it involves a considerable investment and high uncertainty [4] [5].

The discovery of new sources of hydrocarbons during the last decades has not been able to replace the current reserves [6] [7]. According to Alvarado and colleagues, most of the current oil production (at least 50%) comes from mature reservoirs; due to the decline of major new field discoveries, the most likely place to find more oil is the places where it has already been found [6].

Other authors estimate that the proportion of wells producing oil by a natural flow can be as little as 1 in every 20, meaning that most of the global production happens by means of an artificial lift procedure. [8]. Additional recovery methods are known as Enhanced Oil Recovery (EOR). EOR techniques happen during the secondary and tertiary stage of recovery. Injecting a fluid into an operating field represents an extra cost, but it is estimated that it is a profitable process when the price per oil barrel is above $20USD.

The oil industry is currently facing a new era where technology has become an essential part of it; due to the development of computational tools in the last few decades, the advances have been very successful by their incorporation [9].

1.2 Motivation

Being able to predict how production will perform due to different injection strategies is very important, since the best or optimal decision can be applied [10]. Normally, two different optimisation strategies are followed, one for long term production known as reservoir management and one for the daily operations known as reservoir optimisation [11]. Forecasts in the reservoir optimisation strategy have to be more accurate than the ones in reservoir management since immediate actions are required.

Forecasting a possible scenario is feasible by using mathematical models, which can predict how the system will perform under different conditions. These predictions must be accurate and fast to obtain, this is where the main challenges of system identification and reservoir modelling as it is called in the oil and gas industry come. Current reservoir simulation tools rely on geological properties and experimental tests, the accuracy of these software models depends on the quality of the input parameters as well as the experience of the user [12]. It is a time consuming and demanding process, which not always provides accurate predictions.

On the last decade, several techniques have been developed for characterising reservoirs by using historical production data. The idea is that the flow on every well is affected by the flow on other wells from the same field. There are some limitations when characterising a reservoir by only using injection and production data [13]. The reservoir must be under injection and the rates must vary. If new wells are opened or transformed (from/to injection from/to production) their mathematical description becomes challenging, a

relatively big amount of data must be available. the resulting model performance and its quality is highly related to the performance of the model. Even when these limitations exist, it results easier to have computer data where tests can be performed than making real tests on an oil field, where costs increase and failures with dangerous consequences are likely to happen.

For example, there is a technique for finding the relation between an injection well and a producer one by injecting a colourful substance into the reservoir, the amount of colour collected at a production well tells the traceability between a pair of wells. This process does not only represent an extra investment, but it might take a long period of time before obtaining a result, years in some cases. It is even possibe that the analysed wells are not related at all [14].

Belkis and colleagues, performed an experiment using tracer substances in order to find the inter-well connectivity on an operational field. They found it normally takes 4 years to get useful results. On the other hand, tracer recovery rates are around 9%, this complicates the process of finding the relationship between wells. Simulating the reservoir through a computer model can predict this behaviour relatively fast since the dynamics of the system are hidden within its previous input/output rates.

The objective of the research studies introduced on this thesis is to use system identification techniques for analysing and forecasting injection and production rates from an operational oil field. This will provide a tool, which will allow predicting future scenarios and optimising the current injection strategy.

The following image describes the faced scenario of a mature reservoir under injection.

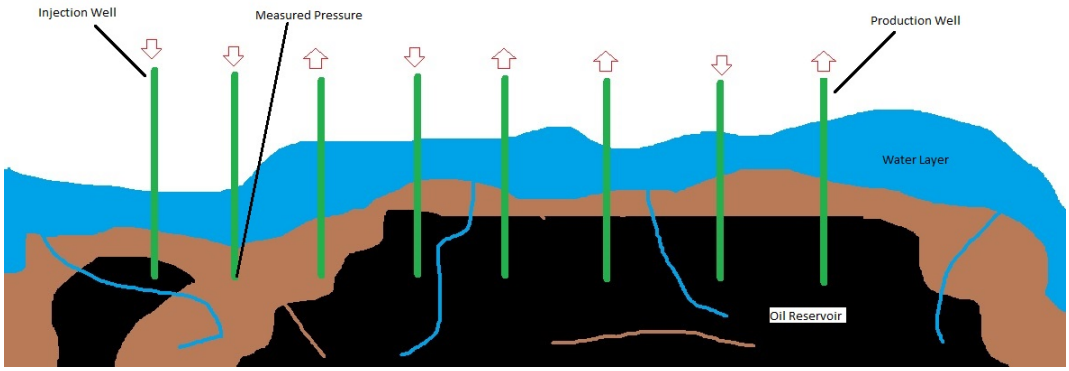

Figure 1.3: Simple Field Under Injection Diagram

From the previous figure, green lines represent drilled wells into the ground. The arrows pointing down indicate the corresponding well is for injection while arrows pointing up indicate the well is for production. From the surface, it results complicated to know how wells are interconnected and how much fluid must be injected.

Understating ground properties and well interconnectivity not only increases production performance but also helps to avoid potential environmental problems, which should be avoided at all times.

1.3 Contributions and Thesis Overview

The development of the research studies shown on this thesis led to significant and satisfactory results. The main contributions and chapter description are summarised as follows:

1.3.1 Main Contributions

The developed algorithms provide an alternative to traditional reservoir simulation methods. The presented methods give an insight into the reservoir's properties which is not possible by using existing data based approaches. According to simulations, the developed tool can significantly improve oil production performance.

- **Polynomial NARX Models for EOR Simulation**
  The application of NARX models for EOR simulation was customised and implemented. The models were capable of capturing the dynamics of an operating field in the North Sea. These models explicitly show which injection wells are the most significant and how long it takes for their effects to be observed on production. This is a cross disciplinary work, making use of system identification algorithms into petroleum engineering.

- **Pruned Multi-Scale Radial Basis Function (MSRBF) NARX Models for EOR Simulation**
  In literature MSRBF NARX models have proven to be effective for system identification. However, the systems on which they have been tested have a low number of inputs. Dealing with a high number of inputs has the risk of selecting model terms with low significance on

the output which would then lead to poor model performance. In order to address this problem, a novel methodology on which several models are estimated by randomly disabling inputs on the system was implemented. A final prediction is estimated by combining the prediction from simpler models using an ensemble. This approach has proven to be very effective for EOR simulation where the number of inputs is normally high.

- **Customised Data Handling for EOR Analysis**
  Historical data coming from an operating field is unique in many aspects. Functioning wells are very likely to be closed for some time due to maintenance or change their role from/to production/injection. This nature results in a dataset with missing values and abrupt changes. An oil field is not a system where the user has control on the past input values. On this thesis, the use of customised methods that deal with the mentioned challenges is demonstrated, this enables the possibility of conducting further analysis. An effective methodology is introduced for finding an appropriate number of iterations when running a Monte Carlo simulation for risk analysis.

- **Risk Analysis Integration with Production Optimisation**
  The concept of risk and the developed methodology in financial mathematics provide an estimation of the likeliness of a certain scenario to happen. The generated NARX models enable the possibility to see how diverse operation conditions will affect oil production. A novel contribution was introduced by designing a future injection strategy using an heuristic optimisation approach known as GRASP applied on the NARX models. With this approach, the operating oil company can

either chose a production goal and evaluate its feasibility or choose a feasibility probability and then design a production strategy which is very likely to be met.

1.3.2 Chapter Description

- Chapter 2: On this chapter, a literature review from the traditional oil production estimation techniques is presented. An EOR modelling case study using multi-layer neural networks is shown as a reference for further comparison.

- Chapter 3: This chapter begins by presenting the Non Linear Autoregressive Moving Average with Exogenous Input Model (NARMAX) philosophy for system identification. The chapter shows how to implement the NARMAX methodology for the estimation of polynomial NARX models based on injection/production data from an operating field. A description of how to pre-process data, validate and select the best model is given. It is shown how to interpret the resulting model equations and link them to the reservoir's physical properties. The problem of characterising EOR production can lead to a high number of variables which would require high computational power. The chapter explains how to gradually increase the number of lags and polynomial degree, this approach leads to feasible model estimation.

- Chapter 4: This chapter begins by presenting MSRBF models. It is shown why applying the MSRBF NARMAX methodology and estimating the models off the shelf would lead to poor results. The contribution of this chapter is the implementation of a novel pruning method that randomly disables inputs so that simpler model structures can be

estimated. The resulting simple models are used together as an ensemble to produce reliable forecasts. The novel methodology reduces uncertainty on the predictions while enabling the use of a high number of system inputs (as required on EOR).

- Chapter 5: This chapter explains the concept of risk and the methodology used for analysing it. The originality presented on this chapter consists in the extension of quantitative and qualitative risk analysis tools usually implemented in financial mathematics to the developed NARX models from the previous chapters. The chapter explains how the results should be interpreted and their impact when forecasting future scenarios.

- Chapter 6: This chapter begins with a review of the main optimisation methods placing special attention on heuristic methodologies and their advantages. The idea is to design a future injection strategy that would maximise oil production. The contribution of the chapter is applying a GRASP algorithm for estimating optimal future injection values. The GRASP optimisation technique is widely used in operation research but has never been used for the problem of designing future injection values. The methodology is implemented on the developed NARX models in conjunction with risk analysis tools from chapter 6. This integration is very important, it allows the estimation of a realistic injection strategy which is very likely to increase oil production. Without this integration, infeasible solutions would be estimated.

- Chapter 7: This chapter provides conclusions for the thesis as well as a description of future work ideas.

1.3.3 Published Articles

- Peñuelas-Alvarez, J. A, Wei, Hua-Liang, and Ibargüengoytia-González, P. (2015). Understanding Water Injection From a Control Engineering Approach Using the ERR-OLS Algorithm. In *SPE Latin American and Caribbean Petroleum Engineering Conference*, Quito, Ecuador.

- Peñuelas-Alvarez, J. A, Wei, Hua-Liang, and Ibargüengoytia-González, P. (2016). Mature Field Data Analysis Using a Polynomial NARX Model. A Case Study From the North Sea. In *Innovation Match MX 1er Foro Internacional de Talento Mexicano*, Guadalajara, Mexico.

- Peñuelas-Alvarez, J. A, Wei, Hua-Liang, and Ibargüengoytia-González, P. (2016). A Novel Single Layer Neural Network Ensemble Model for History Matching. *Journal of Engineering Applications of Artificial Intelligence.* (Under Review)

- Peñuelas-Alvarez, J. A, Wei, Hua-Liang, and Ibargüengoytia-González, P. (2016a). Análisis de Datos para Modelado de Yacimientos y Optimización de la Producción Utilizando Estrategia GRASP. In VIII Congresso Anual Conjunto de Asociaciones del Sector Energético XVIII Congreso Anual de la AMEE, Acapulco, Mexico. **This paper was awarded the first prize from 78 presented articles.**

Chapter 2

State of the Art in Reservoir Simulation

Simulation is the ability to obtain a system's response by using a mathematical description of it (mathematical model). Future input values and conditions can be tested by using a model without the need to actually implement a physical experiment where undesirable scenarios may occur. The accuracy and performance of the simulations depend on the quality of the mathematical model. Since models are only an approximation of the real system, they will always differ from reality.

Almost since the beginning of the oil production era, being able to estimate how much oil can be extracted from the ground has been an important matter. Fast approximations were developed for this purpose: material balance equations, fractional flow curve methods, sweep efficiency estimations [15],[16]. As reservoirs became more complex and higher precision was required, these approximations were not longer sufficient to satisfy the indus-

try's needs. To do so, complex methods based on physical equations were developed. The physical principles that describe fluid flow within the reservoir are: conservation of mass, isothermal fluid phase behaviour and Darcy approximation of fluid flow through porous media [17].

In order to implement these methods, computers are required since the number of equations to be solved is simply too large to be manually executed. Specific software that integrates these principles has been created by diverse companies. The software packages are known as reservoir simulators, the list of the most popular includes: BOAST,CMG, ECLIPSE, Tempest MORE, ExcSim, Nexus, ResAssure, ReservoirGrail and Merlin [15].

There are some drawbacks about using reservoir simulation software, its usage requires specific knowledge about the reservoir's physical properties, if these parameters are incorrect then the software's estimates will also be incorrect. The usage of the software packages is not straight forward and their licence can be considerably expensive. Most ranging 1 million USD.

With the development of modern computers, it is now possible to analyse data by using complex and iterative algorithms. This is a new approach which has not fully been explored. These methods known as machine learning or artificial intelligence have demonstrated their capability for building models that can replicate a system's dynamics. Their application in the oil industry is described as a "technology at an infancy stage" by experts in the field [8].

In theory, most of modern machine learning algorithms could be applied

for estimating future oil production by means of water/gas injection. But the problem of characterising a reservoir is not only about fitting a function which will track certain values. It is also important to physically understand why the field is behaving in a certain way, this enables the possibility of taking the best decisions that will lead to extracting our natural reserves in the most efficient way.

This is the reason for which specific mathematical models have been developed for history matching (this is how training models is called in petroleum engineering), where the model parameters can be linked and interpreted rather than a simple black box prediction model. The reviewed methods presented on this chapter were developed and adapted specifically for this purpose.

2.1 Traditional Methods

2.1.1 Analysis of Decline Curves

This method enables the possibility to estimate how oil production will change over time, it is one of the most popular forecasting methods in the industry. It was developed in 1945 by J.J Alps, it is based on older techniques from the beginning of the 20th century, including the work from Arnold and Anderson (1908), Cutlen (1924), H.N Mash (1928) and Allen (1931) [18].

When the model was first introduced, it was used to assess the oil demand the war required. At the time more wells had to be opened, the existing production wells just could not keep up with the desired production targets any more. The development of this model sets a historical moment when it was

first realised that the trend of extracting oil at a high rate would eventually decay. Before this time, most of the new field discoveries had relatively steady production values.

Arps found that future production values depended on the previous production rates and it was therefore possible to estimate future values by extrapolating the production trend. After analysing data logs from a large number of wells, it was found that the best results were obtained by using an exponential function [18] :

$$P = At^m e^{-Bt} \tag{2.1}$$

Where A is the initial production value, m is the sampling time and B is the decay rate. To find the equation's parameters the data records have to be written on a table, from which the decay rate can be computed as:

$$B = \frac{\Delta P}{\Delta t} \tag{2.2}$$

There are few variants of the method where the decay rate can be recursively re-estimated providing a better approximation. At the time the method was created, the computation capabilities were limited, so a graphical implementation was developed as a complementary tool. By plotting the decline curves in a log scale, the production curves can easily be extrapolated by using a straight line. This approach makes the method very easy to implement.

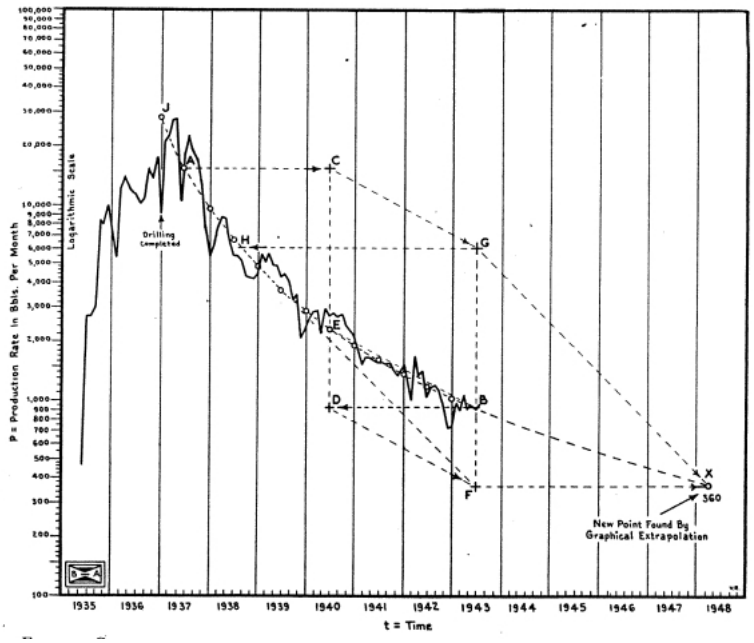

Figure 2.1: Exponential Decline Curve Analysis Example [18]

The previous figure shows how future production values can be extrapolated by using a straight approximation line. The selected points on the curve are used for estimating the decay rate as presented on Equation 2.2. The decay rate represents the slope on the extrapolation line. Modern variants of the decline curve analysis have been developed, for example bootstrap methodology for estimating different decay curves and then forecasting future production by using a final decay rate [19].

As a conclusion, it can be said that the analysis of decline curves provides a simple but effective approximation for future production values. No extensive computations are required. However, the method cannot estimate

what-if scenarios and only considers a decay in production which is not always the case, for example if the field is under fluid injection, oil production can remain steady or even increase. Its application for complex scenarios is therefore very limited.

2.1.2 Material Balance Equation

The material balance equation describes the simplest form of conservation of mass in a reservoir. The model describes how fluids flow within the reservoir due to rock properties. The model was first introduced by Schilthuis in 1936 and has had different adaptations since then. According to the U.S Bureau of Mines, the equation marks the beginning of petroleum engineering as a discipline [20].

The principle of material conservation is given by the following expression:

$$\text{Remaining Fluids} = \text{Initial Volume of Fluids} - \text{Produced Fluids} \tag{2.3}$$

Reservoir analysis through the material balance equation assumes the following behaviour of formation volume factors.

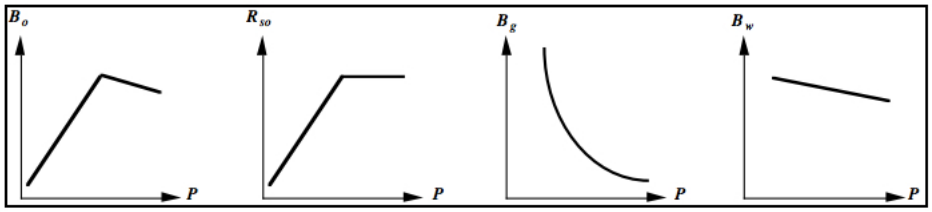

Figure 2.2: Oil, Gas and Water Compressibility Ratios

[21]

These factors show how much oil, gas and water is compressed at a given pressure, where:

$$B_0 = \text{Reservoir oil volume/Standard oil volume} \tag{2.4}$$

$$R_{SO} = \text{Solution gas-oil ratio, vol gas/vol oil} \tag{2.5}$$

$$B_g = \text{Reservoir gas volume/Standard gas volume} \tag{2.6}$$

$$Bw = \text{Reservoir volume of water/Standard water volume} \tag{2.7}$$

$$P = \text{Bottom Hole Pressure} \tag{2.8}$$

The model provides an estimation of the reservoir as a whole, a description for individual wells is not possible. The material balance equation considers the reservoir as a tank, where the fluid pressure and properties are the same regardless of the location within the reservoir. For example, at any production time, it is likely that pressure will drop but the reservoir volume remains the same, the "gap" left by the extracted oil would be filled either by gas expansion or any of the fluids flowing into the reservoir. The general form of the material balance equation for reservoirs is given by:

Withdrawal - Injection = Hydrocarbon Fluid Expansion + Water Influx (2.9)

Before going into the details of the equation, the following terms must be defined as:

$$N = V\phi(\frac{1 - S_{wc}}{B_{oi}}) \tag{2.10}$$

Where N is the initial oil in place, V is the reservoir volume, ϕ is the rock porosity, S_{wc} is the water saturation and B_{oi} is the original volume of oil and dissolved gas. N_p is the cumulative oil production.

R_p is the cumulative gas oil ratio, given by:

$$R_p = \frac{CumulativeGasProduction(scf)}{CumulativeOilProduction(stb)} \tag{2.11}$$

The material balance equation considers the effect of oil expansion with dissolved gas. The difference in volume due to a drop in pressure is given by:

$$B_o - B_{oi} \tag{2.12}$$

Where B_{oi} is the oil volume after a pressure drop, B_o is the original oil volume. The oil mixture is in equilibrium with a gascap, as a result of the reduction in pressure some of the gas dissolved in the oil will be liberated. The volume of the liberated gas is given by:

$$R_{si} - R_s \tag{2.13}$$

Where R_{si} is the total amount of gas in the solution, while R_s is the remaining dissolved gas in the solution after the pressure drop. As a consequence of the extraction of fluids in the form of production the size of the gascap increases, it is defined as follows:

$$mNB_{oi}(\frac{B_g}{B_{gi}} - 1) \tag{2.14}$$

Where B_{gi} and B_g are the initial and final gascap size, respectively. The effect of pore volume reduction is given by the following expression:

$$(1+m)NB_{oi}(\frac{c_w S_{wc} + c_f}{1 - S_{wc}})\Delta p \tag{2.15}$$

Where m is the initial gascap size, c_w is the water compressibility, c_f is the pore compressibility, S_{wc} is the water saturation and Δp is the change in pressure. Now that the terms from the general equation have been defined, the final version of the equation is written as:

$$\begin{aligned} &N_p\left[B_o + B_g(R_p - R_s)\right] \\ &= NB_{oi}\left[\frac{(B_o - B_{oi}) + (R_{si} - R_s)B_g}{B_{oi}} + m(\frac{B_g}{B_{gi}} - 1) + (1+m)(\frac{c_w S_{wc} + c_f}{1 - S_{wc}})\Delta p\right] \\ &+ (W_e - W_p)B_w \end{aligned} \tag{2.16}$$

Where the final term represents the net water influx into the reservoir. W_e is the cumulative water influx into the reservoir, W_p is the cumulative water production and B_w is the net water formation volume factor.

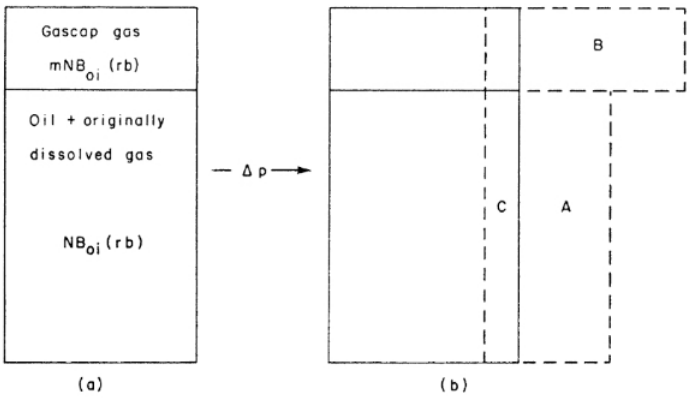

Figure 2.3: Fluid Expansion Diagram

[22]

The block on the left-hand side represents the initial reservoir state with a finite gas cap, where the original fluid volume is given by the solid line. On the right-hand side of the figure, the effect of reducing pressure is presented, as it can be observed, as a consequence of the pressure drop, fluids within the reservoir expand. The pressure drop is a direct consequence of extracting hydrocarbons from the ground as production. In order to estimate future production values, the material balance model has to be evaluated at the time where the pressure drop will occur Δt, this process can be completed in a stepwise way.

The material balance equation can be thought as an extended version of the compressibility definition. The following equation represents the change of volume due to a pressure change.

$$dv = cV\Delta p \tag{2.17}$$

Where c is a constant value indicating the compressibility of a certain fluid, V is the container's volume while Δp is the change of pressure [22].

The method requires the value of certain reservoir parameters which brings uncertainties to the computed estimates. The required values are: size of initial gascap, initial amount of oil in the reservoir, the influx of the aquifer and the reservoir's pressure which in reality is not homogeneous. Even when it is possible to estimate these values, it is very hard to obtain an accurate approximation, which would lead to a poor model performance. In order to reduce the uncertainty on the estimates, it is required to measure all the fluids going in and out every well accurately.

The material balance equation is a methodology that provides a rough approximation of the future production values. In practice, it is very difficult to measure fluid flows and pressure within the reservoir. The method however does provide a good description of how oil, gas and water from different sources interact when the reservoir is under production.

2.2 Capacitive Resistive Model

The idea of the capacitance model begins as an alternative to conventional methods for characterising an oil field; this model was developed at the University of Texas. This model considers the reservoir as an electrical circuit composed of a resistor and a capacitor, where the injection rates simulate a voltage generating a current (flow in the reservoir).

The goal is to use well rates as input and output values to analyse the connectivity between an injector and a producer well pair without resulting into a computationally expensive method [23]. Interwell connectivity refers to how a pair of wells are related, if the reservoir is homogeneous then the interwell connectivity only depends on the location of each well, for this case those wells that are close to each other will be highly interconnected; on the other hand, if the reservoir has different permeability zones interwell connectivity varies in a more complex manner [23].

If bottom-hole pressure records are available, the model provides a better description for the field, this extended version of the Capacitive Resistive Model (CRM). There is another variant of the CRM model called Capacitive Resistive Model Producer-Producer (CRM-P) where the change in fluid flow between producers is considered [24].

The CRM model has proven to be a useful tool for production, optimisation, and performance evaluation of a water flooding strategy at an acceptable level of accuracy [25].

The model is tuned by using a linear regression from the continuity equation. The estimated parameters are a time constant, fraction of injection and a productivity index [23]. The idea is to select those term values which minimise the error between the real measurement and the model's output. The model's equation is:

$$q_j(t_n) = \sum_{i=1}^{N_i} q_{ji}(t_0) e^{-\left(\frac{\Delta t}{t_{ji}}\right)} + \sum_{i=1}^{N_i} \sum_{k=1}^{n} \left[\left(1 - e^{-\left(\frac{\Delta t}{t_{ji}}\right)}\right) \left(f_{ij} I_i^k - J_{ij} \tau_{ij} \frac{\Delta p_{wf}^k}{\Delta t_k}\right) e^{-\left(\frac{\Delta t}{t_{ji}}\right)} \right] \tag{2.18}$$

The CRM model equation has three parameters, a linear gain f, a time constant t and a productivity index J. The term Δp_{wf} represents the change in bottom-hole pressure between samples k and $k-1$. If the pressure values are not known, this term can be neglected assuming there is homogeneous pressure along the reservoir. For every j production well, there are Ni injection wells and n time series samples [23].

The model's equation is composed of three elements. From left to right, the first one is primary depletion, injection input signal and changes in pressure in the producer well.

The model assumes certain conditions, and is not always optimal. For example, the term related to natural decay in production assumes the decay is exponential and this might not always be the case since reservoir dynamics are far more complex than an exponential behaviour. The f connectivity index constant shows the effect an injection well i has on a production well j, the time constant t tells how long it takes for the injection to produce a change in production. The model has been tested on different artificial scenarios produced by reservoir simulators. On these simplified experiments if changes in pressure are discarded the model produced forecasts with a square

correlation of 0.99, when pressure changes are included this index dropped to 0.728 [13].

The CRM model requires injection/producer wells to remain active through all the dataset. However, in reality this is impossible to have since wells are always shut down for maintenance, converted or simply closed. A new variant of the CRM model was introduced by Kaviani where virtual wells are used in order to compensate for this issue. The virtual injectors have to be manually manipulated . As a result the continuity equation might not be convex which leads to the problem of finding appropriate parameter values. As a result, the model's performance and complexity are compromised when analysing a real scenario. Some examples con be found on [13], [25], [26].

In conclusion, it can be said that the Capacitive Resistive Model is a good alternative to traditional reservoir modelling. It is based on an intuitive equation which can be easily linked to the physical phenomena. However, its application on a real dataset is not straightforward, it might not be flexible enough to fit the complex dynamics of the reservoir and might not produce accurate results. The estimation of the parameters is also a non-trivial task since a non-linear optimisation method has to be used for every continuity equation.

2.3 Statistical Reservoir Analysis

Ian Main and his colleagues developed the Statistical Reservoir Analysis (SRA) at the University of Edinburgh. This method offers a similar analysis to the one obtained by the CRM. It can compute inter well communications;

detect flow channels, flow barriers in a field as well as provide a 3-month reliable forecast [27].

According to Main the SRA methodology requires 3 years of historical production/injection data for model estimation [28]. The required variables are flow and pressure rates for every individual well. The model is estimated in the following manner: First, square correlations between pairs of injectors/producers are computed, if a pair has a low correlation it is discarded for further analysis. The search for correlations for a single producer is stopped when multivariate regression coefficient reaches a value of $R = 0.9$. It normally takes between 5 to 25 injection wells to reach a value of $R = 0.9$.

The Statistical Reservoir Model is given by the following equation:

$$\hat{Y}_t = R_k X_{t-k} \tag{2.19}$$

Where $\hat{Y}_t$ is a vector with the forecast rates for all producing wells N while R_k is a matrix with the regression of the parameters. The parameters are composed of the linear terms composed of the selected injection and producing wells by the Bayesian correlation as well as the lagged terms from the current well under analysis, the model can be viewed as a linear ARX model [29].

The model terms are tuned by minimising the prediction error, which is given by:

$$e = \sum_{t=2}^{T} \sum_{i=1}^{N} (y_{i,t} - \hat{y}_{i,t})^2 \tag{2.20}$$

N represents the number of producers and t is the sampling time. It is important to note that the model can also include lagged variables on the correlation matrix R_k, this matrix would be a three-dimensional matrix with array elements: $r_{i,j,k}$: $i = 1,N;j = 1,N + M;k = 1,K$ [30].

The SRA model was tested on a real dataset from the Gullfanks field in Norway. It was found that most of correlated wells lie within a 2km range. The analysed data has flow rates for 106 wells for a period of 11 years, flow rates values were recorded every month. The SRA model was tested as a forecast tool using flow rates from the Gullfanks field. It was found that for a 3-month production forecast the predictions lied within a 95% confidence limit.

In conclusion, the SRA has proven to be an effective tool for analysing production/injection data from an operating field. It explicitly indicates which injection wells are the most significant towards a certain production well. Its extension to real applications is backed by funding from industrial partners for its development . In comparison with the CRM model, it can easily handle incomplete records and shut down periods on certain wells, where the correlation index is penalised. The drawback from the SRA model is its limited long term forecast capability. A 3 month forecast is good for planning a short term strategy, but most oil companies need at least one year of reliable predictions in order to plan their future operations. Further details about the model remain confidential due to an industrial patent.

2.4 Linear Data-Driven Models

Linear models have become the traditional option for analysing and modelling time series [31]. Since the development of the autoregressive model a well-established methodology for fitting its parameters was introduced by Box and Jenkins [32].

The simplest linear model that can represent an input-output relation is a linear difference equation:

$$y(t)+a_1y(t-1)+...+a_{na}y(t-n_a) = b_1u(t-1)+...+b_{nb}u(t-n_b)+e(t) \quad (2.21)$$

Where, $e(t)$ represents white noise error. The adjustable parameters are:

$$\theta = [a_1, a_2, ..., a_{na}, b_1, b_2, ..., b_{nb}]^T \quad (2.22)$$

Equation 2.15 is known as the Autoregressive With Extra Inputs Model (ARX). The autoregressive part is represented by the $y(t-n_a)$ terms, while $u(t-n_u)$ are the extra terms also known as exogenous.
If a new vector is introduced as:

$$\phi = [-y(t-1), ..., y(t-n_a), u(t-1), ..., u(t-n_b)]^T \quad (2.23)$$

The ARX model is given by:

$$\hat{y}(t|\theta) = \theta^T\phi(t) + \upsilon(t) = \phi^T(t)\theta + e(t) \quad (2.24)$$

Where, $e(t)$ is can be approximated to an offset value.

The previous form is the most popular for practical applications since it is intuitive. There are more complex structures but most of them can only be used in certain cases. A good review of these structures can be found on [33].

In order to estimate the value of the model's parameters given by vector θ, the most popular methods are the least square estimator and the maximum likelihood estimator, details of how these estimators are obtained can be found at [33]. The Least Squares Estimator is defined as:

$$\theta = \left[\frac{1}{N}\sum_{t=1}^{N}\phi(t)\phi^T(t)\right]^{-1}\frac{1}{N}\sum_{t=1}^{N}\phi(t)\phi^T(t)y(t) \tag{2.25}$$

N is the number of samples and ϕ is defined as in Equation 2.17. The Maximum Likelihood Estimator is defined as:

$$\theta(y^N) = \frac{1}{\sum_{i=1}^{N}(\frac{1}{\lambda_i})}\sum_{i=1}^{N}\frac{y(i)}{\lambda_i} \tag{2.26}$$

N is the number of samples and λ is the standard deviation within the data samples. In this case y is not the output but the corresponding regressor variable from matrix ϕ.

In literature, there are some examples where ARX models have been fitted for modelling oil production using water injection rates as inputs. For example, Van Essen and colleagues, used an ARX model to track production data. The data was artificially obtained by a reservoir simulator, as a result a 12th order model was able to capture the system's dynamics [34] .

Marte implemented an ARX model for fitting data from three different reservoirs in the North Sea, as a conclusion of the studies, it was found that

a radial basis function could approximate the data better than the linear model[11].

Kun-Han used a multivariate ARX model to characterise a reservoir [35]. As a conclusion of the studies it was determined that ARX models should only be used to analyse the dynamics of the system by observing the maximum lags on the final equation rather than a forecasting tool. One of the advantages of ARX models is that they can easily be retrained using a time window, this is convenient in the case certain wells are shut down. The model would then easily recapture the new dynamics on the system. On the other hand, the model's prediction horizon would not be very large since the model would only be valid for a short period of time.

According to the literature review, linear models should only be used as an extra tool for analysing the data and not for reservoir modelling. It is well known that reservoir dynamics are non-linear, therefore a linear model is not sufficient for replicating the dynamics.

2.5 Multi-Layer Neural Networks

Artificial neural networks as the names suggests, try to imitate how real neurons on a brain communicate with each other. The methodology is inspired on a biological system, where the system is trained through previous observations and it's able to replicate the relationship on unobserved data.

Neural networks have gained researcher's attention during the last decades and have been successfully used on diverse applications. A few examples on

the list include: pattern recognition, system identification and classification problems. In the area of reservoir modelling several authors have used neural networks as an alternative analysis tool. For example DeJonge used a neural network approach to find discrepancies from the expected field behaviour and trend [36]. Saputelli and colleagues developed a hierarchy model for determining which were the most critical elements in a producing field using a neural network model. This approach allows the personel to intervene on the most important variables [37]. Esmaili used a neural network for understanding the process of hydraulic fracturing revealing the main elements involved [38]. A good review of the applications of neural networks in the field of reservoir engineering was made by Tahar [39], some of the applications include: prediction of petrophysical parameters in reservoirs, tight gas reservoir development, prediction of natural and induced fracture zones and monitoring of physical variables (porosity, permeability and pressure).

Different network architectures have been developed for specific applications. Independently of the problem, neural networks have shown their advantages: they are easy to train, easy to implement, have great approximation capabilities and offer great relationship between non-linear input-output variables.

An artificial neural network creates input-output relations in the following manner: First, input data coming from every neuron goes into a non-linear function, such as a radial basis function. The output of this function is then multiplied by a constant value known as weight. The result of this operation is then added up with the multiplied outputs from all the other neurons. This result is then the new input data for a neuron on the sub-

sequent layer. The process is repeated through all the hidden layers until a final output value is estimated. Selecting the connection weights between neurons and layers is known as the training process [40].

A multi-layered neural network can be represented by the following diagram.

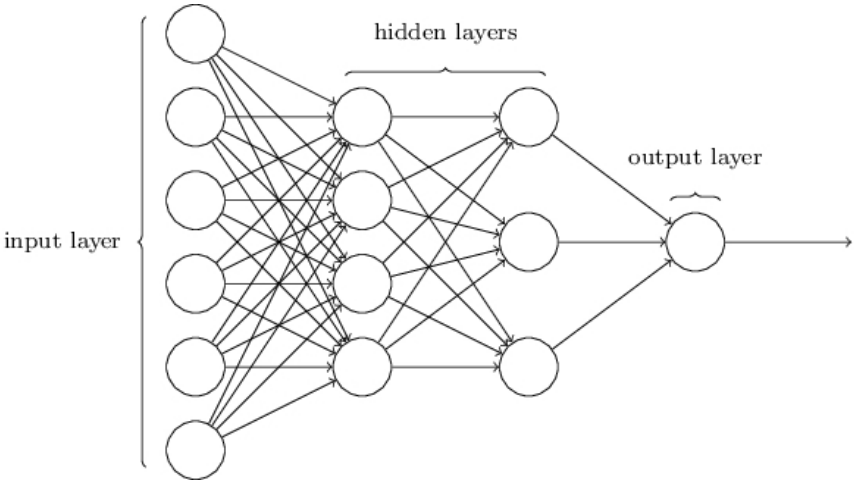

Figure 2.4: Neural Network Structure

[41]

Multi-layer neural networks are probably the first approach one might try for fitting non-linear input-output data. It is a well-known standard method in the areas of machine learning and data mining and has demonstrated its capability finding non-linear input-output relations [41].

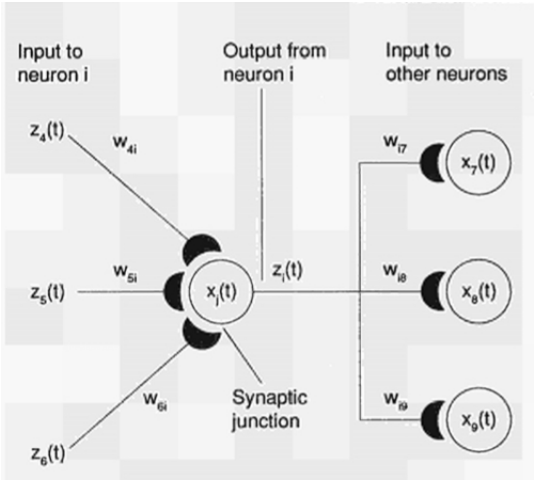

Figure 2.5: Multi-Layer Neural Network Structure

[41]

A feedback architecture is used on this study since the output signal is transmitted back into the input layers. A Gaussian function was used as the activation function on the neurons. The activation function is the element on the network responsible for establishing non-linear relations between the neurons.

There are many choices for the activation function, the most popular ones are radial basis functions. Radial basis functions estimate a value depending on the distance between the function's centre and the input value. The most popular radial basis functions are: Gaussian functions, multi-quadratic functions, inverse multi-quadratic functions, thin plate spline function, cubic function and linear function.

Gaussian functions are one of the most commonly used functions, they

have great generalisation capability and have been well studied. The idea of using radial basis functions is to approximate the observed data by a multi-dimensional surface. A Gaussian function is defined by the following function:

$$f(x) = w(e^{-\frac{1}{2\sigma}(x-c)^2}) \tag{2.27}$$

From equation 2.27, w is a constant scaling factor, σ is the standard deviation also know as scale which determines how wide the function is, or how sensitive the function is with respect to the centre's value c. When training a radial basis function model, these three parameters have to be tuned in order to obtain a good approximation to the dataset.

Training a multi-layer neural network is not a straight forward process. The idea is to find the connection weights, standard deviation and centre values which minimise the error between the network's output and the measured data. Diverse training algorithms have been developed for this task, the list includes Bayesian regularisation, Levenberg-Marquardt and gradient descent algorithms, further details can be found on [42].

For the present study, the Levenberg-Marquardt (LM) algorithm was used for determining the network's parameters. The LM algorithm has proven to be very effective for non-linear least-squares minimisation problems, the solution is found iteratively until a pre-defined threshold value is reached [43] [42]. On the present analysis, the number of maximum iterations (epoch) was set to 100. For most training cases, a solution was found within 10 epochs.

Constructing a neural network with an optimal number of elements has been an important subject of study in the field of machine learning. Most authors conclude that the network elements such as number of lags, training iterations, number of neurons and hidden layers should be determined empirically [44] [45].

Following a design suggestion from Maren [46]. A first guess of number of hidden neurons was determined from the following expression:

$$H = 2N + 1 \tag{2.28}$$

Where H is the number of hidden neurons and N is the number of inputs. To investigate the effect of hidden neurons, a set of models with different hidden neurons were estimated. Only one hidden layer was used, the use of more hidden layers is a subject of current study, it is known as deep learning, it requires high computational power and is normally used for very complex problems like speech recognition and autonomous vehicles [47].

On complex applications each one of the hidden layers can learn or recognise a single feature. For example, if we want to identify red cars on different photos, we would need to identify red objects first and then objects that look like a car. These two features are so different that cannot be modelled by a single function, as the complexity of the problem increases so does the need for additional layers [48].

2.5.1 Scott Field, a Case Study From the North Sea

The analysed dataset comes from Scott, a field in the North Sea. The field started production operations by late 1993. According to the records, on its best month, production peaked 830,000 monthly oil barrels until it declined to 70,000 bbl. The dataset has 69 monthly records (5.75 years). The data is published by the United Kingdom's Department of Energy and Climate Change. The location of the field is displayed on the following figure:

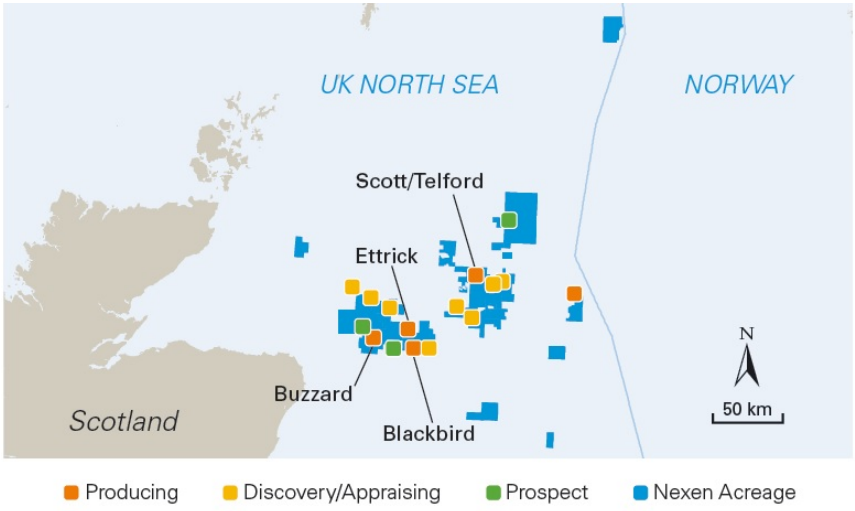

Figure 2.6: Scott Field Location

[49]

The first step was to understand the raw values. The dataset for oil production are monthly records from November 1993 until December 1999, there are 24 production wells. Water injection records are also given monthly from April 1994 until December 1999, there are 20 injection wells.

Since it is assumed the field is a causal system on which the output (Oil

production) purely depends on the previous inputs (Water injection), it results convenient to constrain the analysis from April 1994 until December 1999, eliminating the first 4 months of oil production. As a result, only 69 records from the original 74 were analysed. The records were stored in two matrices, one with a dimension $N \times O$ for oil production rates an another $N \times W$, N being the number of records, while O and W represent the number of oil production and water injection wells respectively.

Oil production and water injection rates are both given in m^3. As a first step to understand the magnitude and behaviour of the system the raw data is plotted as follows.

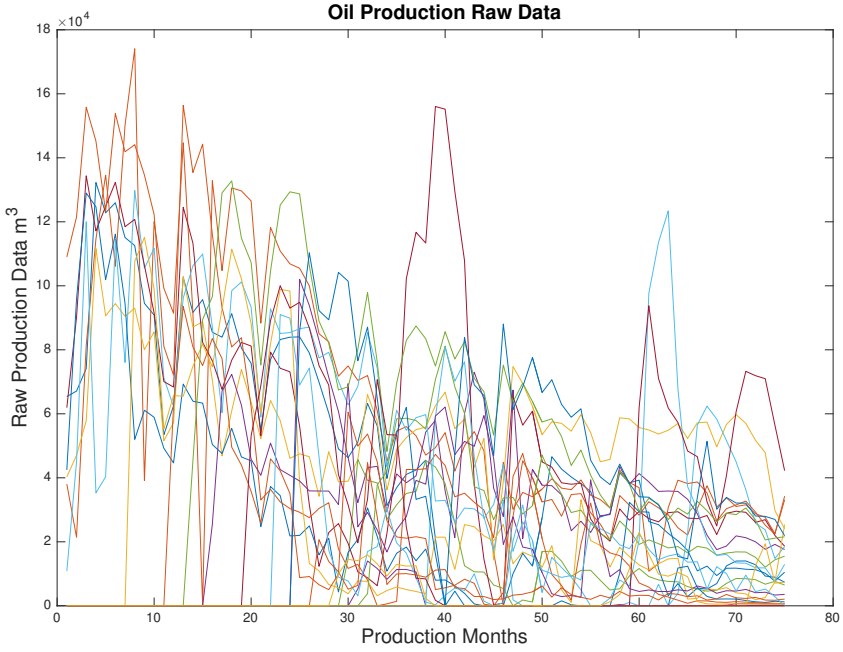

Figure 2.7: Raw Oil Production Data

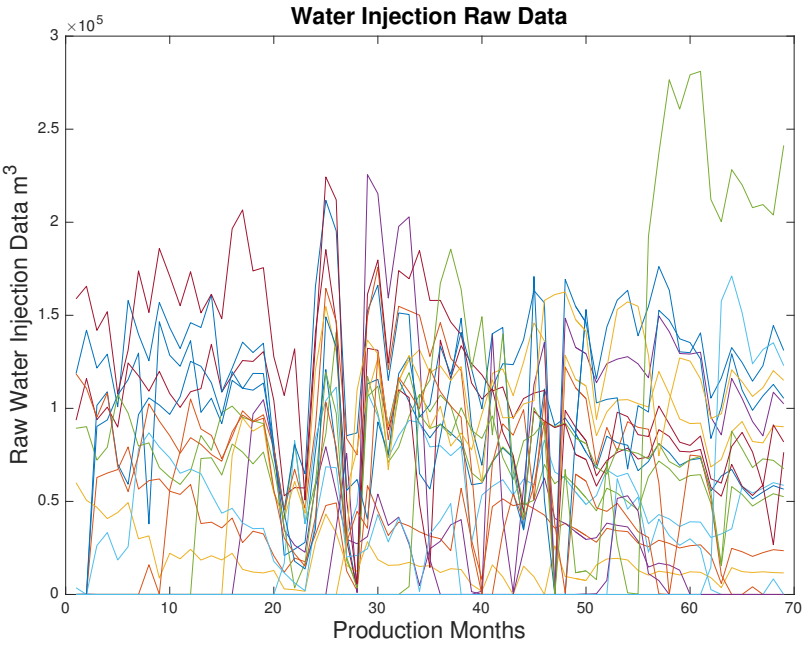

Figure 2.8: Raw Water Injection Data

Each of the lines represent the corresponding oil or water rates from a particular production/injection well. From the previous plots, it results evident that oil production follows a decline trend for most of the operating months as it would normally be expected from a mature field. On the other hand, water injection seems to remain at a constant value. One assumption would be that the constant injection rates are not sufficient to compensate for the production decline, to compensate for this loss an increase in injection should have taken place.

On average the field produces 630,000 m^3 of crude oil a month while 1.25Mm^3 of water are injected into the ground. It can easily be seen that

the total volume of injected water is almost twice of the produced oil. This behaviour is appreciated by the following plots where total oil field production and water injection from the whole field are shown.

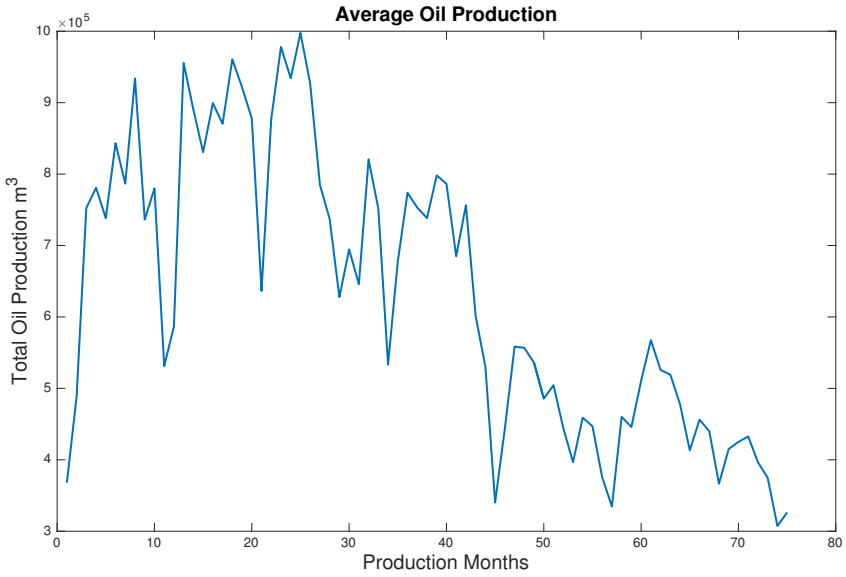

Figure 2.9: Total Oil Production

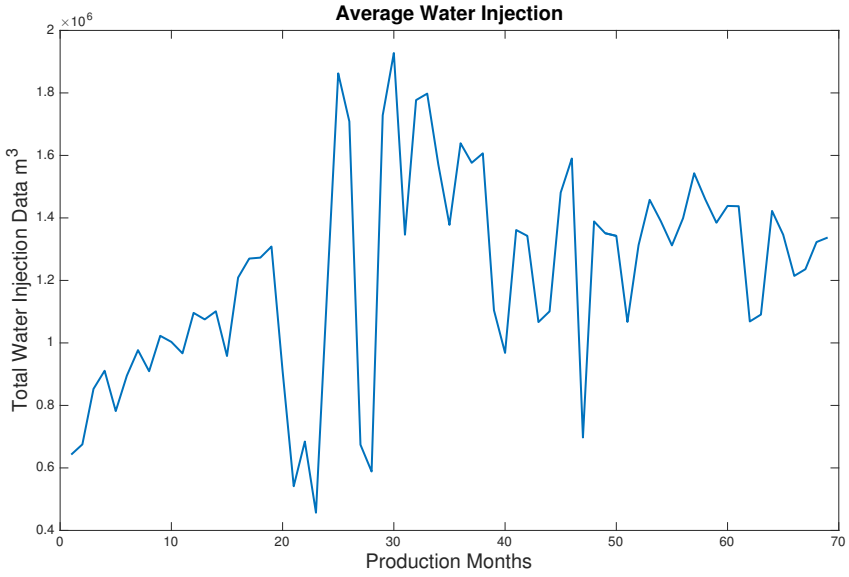

Figure 2.10: Total Water Injection

After having a general understanding of the field, its past behaviour and record values, the next step was to normalise all data values. This is an important step since we want to extract the effect and relation input variables have on the output. Even when the all the raw values are given on the same measurement unit (m^3), it might be possible that a well experiencing small magnitude values has a greater effect on the output compared to other wells with higher values. The normalisation was accomplished using the following expression [50]:

$$x' = \frac{x - min_x}{max_x - min_x} \tag{2.29}$$

Where, x' is the new scaled data point, x is the unscaled raw point while min_x and max_x are the minimum and maximum values from the original

dataset. By using the previous expression all records were scaled from 0 to 1, regardless of their original magnitude.

The following step was to identify any possible outlier measurements on the dataset which might lead to incorrect analysis. A simple but effective procedure was used to identify possible outliers. All data points were plotted against 96% confidence band limits considering there is a Gaussian distribution within the data set, if a point lied outside the limits it would be considered as an outlier suspect. If the point is an outlier, its value would be changed to the average of the 2 neighbour points, this a suggested method by Fortuna [50].

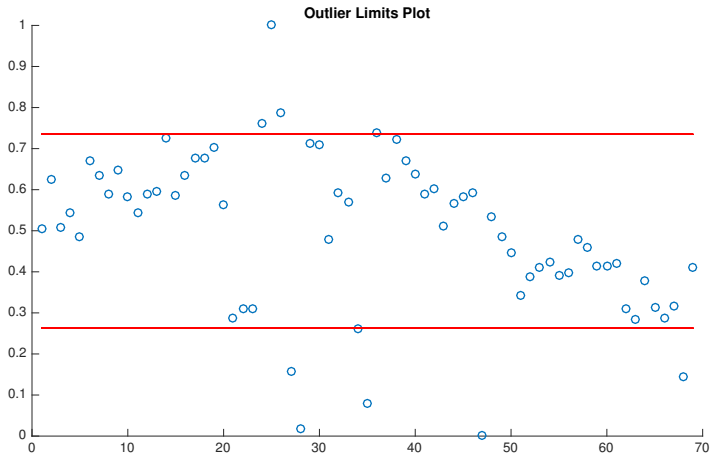

Figure 2.11: Outlier Plot

There is no perfect method for detecting outliers, most selection criteria solutions require interpretation from the user for appropriate selection. However, there is a limit to how fast physical systems can respond to input

changes, and that behaviour must be considered for outlier selection giving us a clue of what could be an incorrect measurement. According to Holdaway [8], outlier detection is always done manually within the oil industry due to the great differences that may occur within different wells. As a result, the process can be tedious and time consuming.

On Figure 2.11 it can be seen that point 26 is very likely to be an outlier since none of its neighbour points are close and an abrupt change is shown. On the other hand, point 36 which is also outside the limits is likely to be a correct measurement since it follows the trend and magnitude experienced by its neighbouring points.

The Scott field dataset contains a few missing values, as most real datasets. In order to estimate a model, it is required to have a dataset without any missing values. There are different methods for the estimation of missing values; Graham provides a good review of the most popular [51].

Since there were not that many missing values on the dataset (at the most 5 on some well records), the approach used on this case study was to fit a linear autoregressive model and estimate according to the previous records what the most likely value for the missing data point would be. This method is straightforward to implement and very effective. Local trends within the records are not lost and the model parameters are fast and simple to obtain. Every well's dataset was treated as an individual time series, from which a linear autoregressive model was determined using a window containing the 7 previous records (Before the missing point). The window size was determined by using a property of the partial autocorrelation function, when the function

crosses the horizontal axis, given by a zero value, it indicates the maximum lag dependency on the time series [32]. The following plot shows the partial autocorrelation function for injection well 5 as an example.

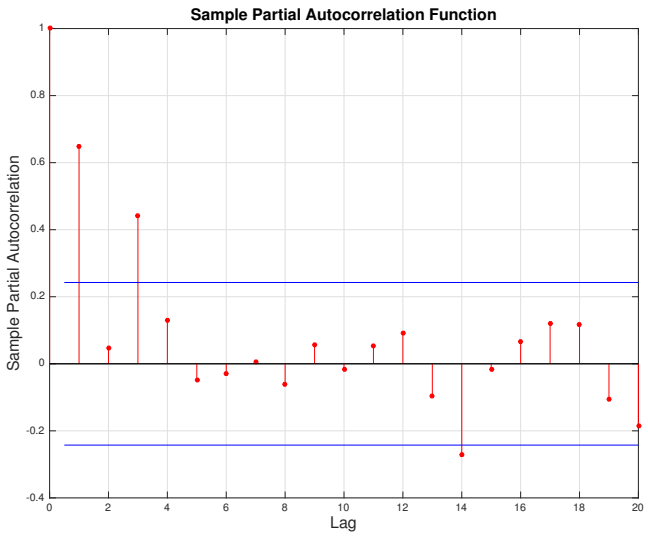

Figure 2.12: Partial Autocorrelation Function

From the previous plot, it results clear that the partial autocorrelation function cuts-off at a lag value of 7. There are different algorithms for estimating the partial autocorrelation function, for the present case study Matlab's built in "parrcor" function was used. A full explanation can be found on the work developed by Box and Jenkins [32].

On the presented experiments, the number of lags was gradually increased from 1 to 12 (representing a year of production) for every set of hidden neurons configuration. Based on these combinations, diverse models were estimated. The models were evaluated using the MSE metric. Data was

split as follows: 57 records for training and 12 for validation, considering that oil companies plan their future operations one year in advance. The performance of the best models can be appreciated on the following table. For full results see Appendix A.

Delays	Hidden Neurons	MSE
12	21	1.25E-03
1	30	1.78E-03
5	40	2.20E-03

Table 2.1: Best Multi-Layer Models

According to Table 2.1 the best model is the model with 21 hidden neurons and 12 delays. The model's response can be appreciated on the following plot.

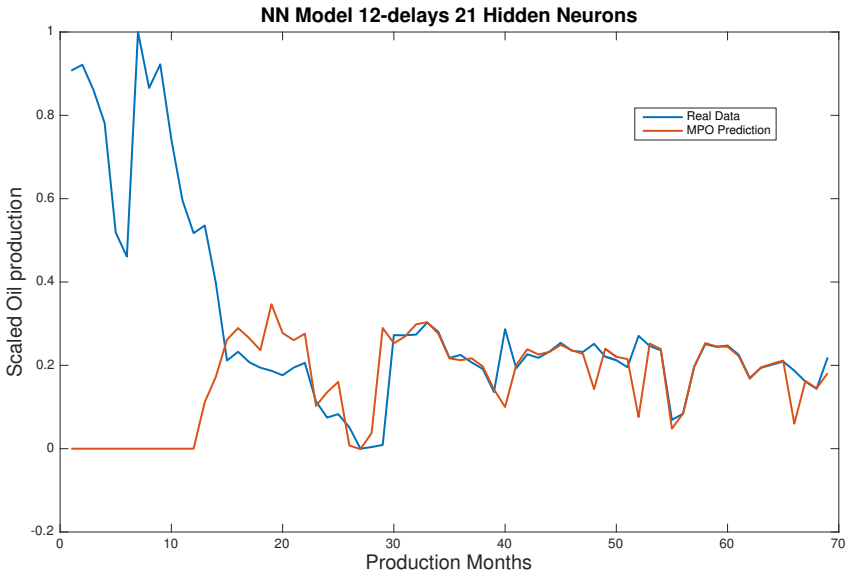

Figure 2.13: 12-Delays 21-Hidden Neurons NN Model Response

The model's structure is as follows:

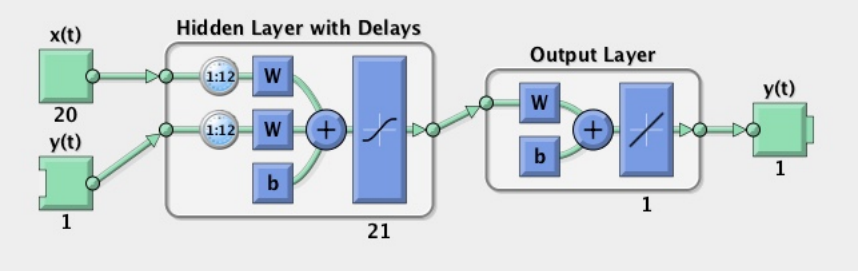

Figure 2.14: 11-Delays 2-Layers NN Model Structure

From the previous results, we can see that training a multi-layer neural network using a standard approach does not produce the best predictions. The resulting model is a black box model, we can't obtain further information about physical properties of the system.

2.6 Conclusions

The current chapter provides a survey of the latest methods for modelling an oil field under production. Traditional methods are very effective providing an explanation how oil, gas and water interact in the reservoir when pressure changes appear as an effect of production. However, their accuracy is limited. In the case of the analysis of decline curves, the method assumes that production will always decay which might not always be the case. Oil production can experience steady values and even increase if the field is subject to fluid injection. On the other hand, the material balance equation can include the effects of injection, but can result in very complex equations which are hard to solve and require accurate measurements of fluids within the reservoir like water flow from an aquifer, these measurements are very

hard to obtain. As a consequence, the model's predictions are likely to be inaccurate.

Data based models offer a more flexible approach. These methods learn from the previous records with the objective of replicating the behaviour of the system. Linear models are useful for revealing some of the field's properties but have failed to provide reliable forecasts since reservoir dynamics are non-linear.

The Capacitive Resistive model is capable of capturing non-linear reservoir dynamics, offers intuitive equations as an analogy to an electrical circuit where physical properties of the field can be revealed. However, when working on a real dataset, solving the equations results a very complex process where non-linear optimisation techniques have to be applied. When dealing with missing values virtual wells have to be introduced, as a result analysing the field using the CRM can be a challenging non-intuitive task.

The Statistical Reservoir Analysis model is capable of revealing inter-well relations and capture non-linear dynamics of the field but can only provide 3 months of reliable forecasts which might not be sufficient.

Multi-layer neural network models are now a standard method when modelling non-linear systems. According to the presented experiments, these models can track oil production with a limited accuracy lacking understanding the physical properties of system under analysis which is fundamental in EOR modelling.

As we can see there is a need for a new reservoir modelling method that can handle all the challenges: flexibility, non-linearity, reveal physical properties and provide reliable long term predictions.

Chapter 3

New Practical Application of the Non Linear Autoregressive with Exogenous Input Model in Petroleum Engineering

3.1 The Non Linear Autoregressive Moving Average with Exogenous Input Model (NARMAX) System Identification Methodology

Data based models offer diverse advantages compared to traditional methods when identifying and modelling complex systems, where the required equations based on physical properties are simply too many or too complex to accurately be tuned to produce reliable estimates. The list of complex systems includes history matching tasks, where reservoir dynamics are non-

linear, unique for every field, time dependent and have multiple inputs.

The task of finding mathematical relations between input and output measurements is known as System Identification in the field of automatic control [52]. The following diagram shows a representation how the analysed variables on the system are related.

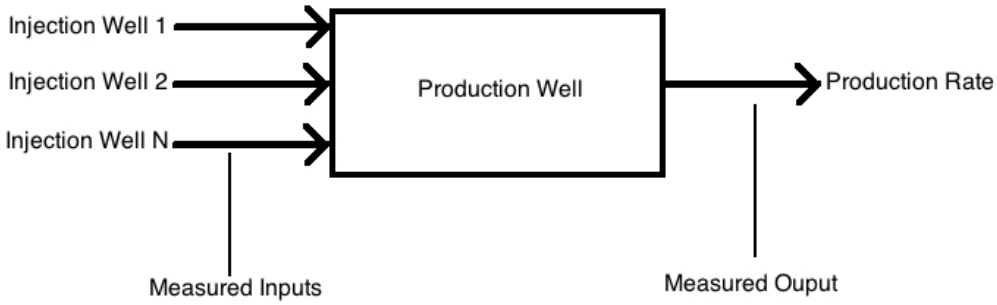

Figure 3.1: MISO System Identification Diagram

As it can be observed on Figure 3.1, to understand the effect of water injection on oil production from an operating field, it is necessary to build a mathematical model for every production well. The mathematical model explicitly indicates the relationship between injection rates to oil production.

Most of non-linear data based models such as: multi-layer neural networks, bayesian networks, fuzzy logic or probabilistic models result in black box structures providing no explicit description of the physical system under study. There is therefore the need for a model that which can adress this issue [53].

To deal with this issue, during the 1980's Billings and colleagues developed a methodology for non-linear system identification. This methodology is flexible since it is based on input-output data measurements and provides transparent equation terms which can be directly linked to the physical parameters from the analysed system [54], [55], [56]. The philosophy is known as NARMAX system identification which stands for "Non-linear Autoregressive with eXogenous input Model".

The generic form of the NARMAX model is given by the following expression [55]:

$$y(k) = F[y(k-1), y(k-2)...y(k-n_y), u(k-d-1)...u(k-d-n_u)+e(k)] \quad (3.1)$$

Where, F is a non-linear function, $y(k)$ is the system's output, $u(k)$ is the system's input while $e(k)$ is additive noise. Since the model is time dependent lagged values must be considered, n_y and n_u are the maximum output and input lags, respectively.

The NARMAX system identification methodology can be summarised on the following steps:

1. Structure selection.
2. Parameter estimation.
3. Model validation.
4. Prediction.
5. Analysis.

The first step consists in choosing a non-linear function F and determining the lags. Afterwards, the estimation of the corresponding terms must take place.

Depending on the selected parameters, a different number of candidate terms will be estimated, in most cases only a few of these terms will be required to form a final model. To select the corresponding terms a powerful methodology has been incorporated into the NARMAX approach, this methodology is known as Forward Orthogonal Least Squares Algorithm (FROLS).

Once the model's terms have been defined. The following step is to verify if the model can properly replicate the system's dynamics. This is achieved through several validity tests. If the model passes the validation tests, it can be used for predicting diverse scenarios as well as understanding the physical properties within the system.

3.1.1 Polynomial NARX Models

From the possible non-linear forms F on Equation 3.1 could take, polynomial structures have proven to be very effective for replicating non-linear data as well as being capable to relate its terms to the physical components and behaviour of the system under study, this gives direct model interpretation. Some of the examples include: characterisation of robot behaviour, identification for space weather and the magnetosphere, tracking iceberg movement in Greenland, understanding of electroencephalography data, analysis of a

Fly vision, modelling of synthetic bioparts, identification of metal-rubber damping devices and forecasting tide levels [57], [58], [59], [60], [61], [62], [63].

A polynomial representation of a NARMAX model is given by the following expression [55]:

$$\begin{aligned} y(k) = \theta_0 + \sum_{i_i}^{n} f_{i1}(x_{i1}(k)) + \sum_{i1=1}^{n}\sum_{i2=i1}^{n} f_{i1i2}(x_{i1}(k), x_{i2}(k)) + ... \\ + \sum_{i1=1}^{n} ... \sum_{il=i_{l-1}}^{n} f_{i_1 i_2 ... i_l}(x_{i_1}(k), x_{i_2}(k), ..., x_{i_l}(k)) + e(k) \end{aligned} \tag{3.2}$$

Where, l is the degree of the polynomial, $\theta_{i_1 i_2 i_m}$ are the linear model parameters while n is the sum of lags given by:

$$n = n_y + n_u + n_e \tag{3.3}$$

When dealing with a multivariate case, the corresponding lag (n_y,n_u or n_e) is the sum of lags from the all the related variables.

Equation 3.2 can also be written in terms of the linear model parameters as:

$$\begin{aligned} y(k) = \theta_0 + \sum_{i_i}^{n} \theta_{i_1 i_2}(x_{i1}(k)) + \sum_{i1=1}^{n}\sum_{i2=i1}^{n} \theta_{i_1 i_2}(x_{i1}(k), x_{i2}(k)) + ... \\ + \sum_{i1=1}^{n} ... \sum_{il=i_{l-1}}^{n} \theta i_2 ... i_l(x_{i_1}(k), x_{i_2}(k) ... x_{i_l}(k)) + e(k) \end{aligned}$$

(3.4)

It is important to know that for a multivariate model, the degree of the polynomial is given by the term with the maximum degree.

The number of terms from a polynomial model is given by:

$$M = \frac{(n+l)!}{n!l!} \tag{3.5}$$

Equation 3.5 shows how the number of candidate terms can be very large. For example a third order, 3-input system with all input lags $n_u = 3$ and single output with $n_y = 2$, would have $M = \frac{(3*3+2+3)!}{11!*3!} = 364$ candidate terms. The NARMAX methodology always looks for parsimonious model structures. As a result, only the terms that significantly contribute to the system's output will be kept at the final model [55].

3.2 The FROLS Algorithm for Term Selection

Estimating a model which can replicate the systems' data without being over-fitted or under-fitted is a subject which has been a subject of study in the field of system identification for the last decades [33], [64], [65]. The number of terms on a model is directly linked to how well the model can fit training and unseen data. In most cases, the selection of a large number of terms would lead to over fitting having poor generalisation performance. On the other hand, selecting a small number of terms will result in poor prediction on unseen data.

In literature, there are broad examples of data fitting by a model when the model's structure is already known. On these cases it is easy to iden-

tify if the model has selected the correct input variables. The complexity of this approach is a lot lower than most real system identification applications where we have no control of the data we gather and need to understand the properties of the system. In such situations, we can't know in advance what are the maximum lag values or the non-linearity degree of the system understudy.

In order to deal with the problem of estimating models with an optimal number of terms, the FROLS [66], [67] [68]. The algorithm selects the most significant terms in their order of contribution towards the output variable. Term selection is done at every iteration step. This approach is convenient since it can be intuitively seen which terms are the most important and to what degree, their contribution is estimated as a percentage of the output variance. The FROLS algorithm can be summarised as follows [55]:

Step 1- Assume there is a number of candidate terms M stored on a dictionary $D = p_1, p_2, ..., p_M$, for $m = 1, 2, ...M$, compute the output variance $\sigma = y^T y$, calculate the contribution of all terms as [55]:

$$g_m^{(1)} = \frac{y^T q_m}{q_m^T q_m} \tag{3.6}$$

$$ERR^{(1)}[m] = (g_m^{(1)})^2 \frac{(q_m^T q_m)}{\sigma} \tag{3.7}$$

$$l_1 = arg \max_{1 \leq m \leq M} \{ERR^{(1)}[m]\} \tag{3.8}$$

Let

$$a_{11} = 1 \tag{3.9}$$

$$q_1 = p_{l_1} \tag{3.10}$$

$$g_1 = g_{l1}^{(1)} \tag{3.11}$$

$$err[1] = ERR^{(1)}[l_1] \tag{3.12}$$

Where g_m is the correlation coefficient between the output and each candidate term m. The ERR coefficient represents how much of the output variance can be explained by candidate term m. Variable l_1 shows which of the candidate terms m has a higher contribution towards the output variance.

At this stage, the most significant term has been selected and removed from dictionary D containing all remaining candidate terms. The term's contribution has been stored at the err variable.

Step 2-(s≥2)

The same methodology from first step is repeated but the remaining vectors (terms) have to be orthogonalised. Therefore, let $m \neq l_1$, $m \neq l_2, ..., m \neq l_{s-1}$. For $m = 1, 2, ..., M$ orthogonalise vectors using the Gram Schmitt algorithm as:

$$q_m^{(s)} = p_m - \sum_{r=1}^{s-1} \frac{p_m^T q_r}{q_r^T q_r} q_r, p_j \epsilon D - D_{m-1} \tag{3.13}$$

$$g_m^{(s)} = \frac{y^T q_m^{(s)}}{(q_m^{(s)})^T q_m^{(s)}} \tag{3.14}$$

$$ERR^{(s)}[m] = (g_m^{(s)})^2 \frac{(q_m^{(s)})^T q_m^{(s)})}{\sigma} \tag{3.15}$$

$$l_s = arg \max_{1 \leq m \leq M} \{ERR^{(s)}[m]\} \tag{3.16}$$

Let

$$q_s = q_{l_s}^{(s)} \tag{3.17}$$

$$g_s = g_{l_s}^{(s)} \tag{3.18}$$

$$a_{r,s} = \frac{q_r^T p_{l_s}}{q_r^T q_r} r = 1, 2, ..., s-1 \tag{3.19}$$

$$a_{ss} = 1 \tag{3.20}$$

$$err[1] = ERR^{(s)}[l_s] \tag{3.21}$$

From equation 3.13, $q_m^{(s)}$ represents the orthogonalised vector at selection step s, p_m is a candidate term from dictionary D, q_r is the selected vector from step $s-1$.

Model terms are selected by using an orthogonal transformation. This transformation comes in the form of $P = WA$, where W is a NxM matrix and A is a MxM upper triangular matrix. From equation 3.14, $g_m^{(s)}$ is an auxiliary parameter vector, this vector is related to the model term parameters Θ as:

$$A\Theta = g \tag{3.22}$$

From equation 3.17, q_s is the selected orthogonalised vector at step s. From equation 3.18 g_s is the auxiliary vector term at step s. From equation 3.19 $a_{r,s}$ is the corresponding element r,s from upper triangular matrix A. The diagonal elements of matrix A are 1, therefore element a_{ss} is one as shown on equation 3.20.

The term selection procedure is stopped when the sum of values within variable $err[s]$ has reached a threshold limit. As a rule of thumb, the threshold value is normally set to 0.95, representing 95% of the output's variance.

The name of this limit is known as Sum of Error Reduction Ratios (SERR) it is given by the following expression:

$$SERR = \sum_{s=1}^{M_0} err[s] \tag{3.23}$$

Where, M_0 is the number of selected terms to form the final model structure. The selection of extra terms can also be manually adjusted for fine model tuning. For example, the procedure could be stopped when the contribution of a term to $err[s]$ is less than a certain threshold regardless of the SERR value.

In most cases the number of selected terms M_0 is much smaller than the number of candidate terms from dictionary D. After all the candidate terms have been selected, the linear parameters have to be estimated. This can be done by using the least squares formula:

$$\theta = (X^T X)^{-1} X^T Y \tag{3.24}$$

Where X is a $N \times M_0$ matrix (N being the dataset length and M_0 the number selected terms). Y is a $N \times 1$ column vector with the output data. Once the corresponding parameters have been estimated, the final model is given by:

$$y(k) = \sum_{i=1}^{M_0} \theta_i q_i(k) + e(k) \tag{3.25}$$

3.2.1 NARX Model Estimation

Motivated by the success and advantages of NARX models (see section 3.1.1 for examples of practical applications), on this thesis their application for

EOR analysis and modelling is presented.

In literature, there are many examples of alternative modelling approaches for data fitting coming from a reservoir simulator, some examples can be found on [69], [70], [71] and [72]. The complexity of most of these examples is low compared to a real operating field, it is on the later cases where the models can play an important role rather than fitting data from a computer simulation. Instead of using a reservoir simulator to generate the injection and production data from a simple field, the NARMAX methodology was tested on a real dataset.

After the data had been briefly analysed and pre-processed, the NARMAX methodology could then be implemented. The polynomial models that were estimated for this case study follow a Multiple-Input Single-Output (MISO) structure. Every production well was considered as a single-output while all the injection wells were considered as its inputs. This means that in order to estimate a model for the whole field, multiple models would have to be estimated (one for every production well).

Estimating a model which would consider all the outputs results inconvenient for multiple reasons. The number of candidate terms would simply be too large, for example using 12 delays for all variables and a third order polynomial would lead to 24,393,776 candidate terms which is excessive for a standard computer to handle. The attractive property of polynomial transparent equations which can easily be linked to physical would not be that obvious due to the high complexity on the final model equation. By using a single output models, the detail and clarity the equations provide is a lot

better for interpretation.

A NARMAX model is composed of three types of terms and their combination: inputs, outputs and noise. For the present case study only a final additive noise term was considered. Including noise terms increases the prediction performance but it results hard to find a relation between the model terms and the field's physical parameters. The estimated models shown on this chapter are therefore polynomial NARX models, discarding complex noise terms from Equation 3.4.

The main objective of history matching is to create a model which can replicate the dynamics of the field under analysis. This enables the possibility to estimate how the field will perform depending on certain input values without having to wait and physically implement these scenarios. When it comes to predicting future scenarios there are two types of estimations, One Step Ahead (OSA) and Model Predictive Output (MPO). The difference of how these estimations are computed is shown using the following model:

$$\hat{y}(k) = y(k-2) + y(k-1) + u(k-1) \tag{3.26}$$

The sequence for OSA predictions should be computed as follows:

$$
\begin{aligned}
\hat{y}(1) &= NA \\
\hat{y}(2) &= NA \\
\hat{y}(3) &= y(1) + y(2) + u(2) \\
\hat{y}(4) &= y(2) + y(3) + u(3) \\
\hat{y}(5) &= y(3) + y(4) + u(4) \\
&... \\
\hat{y}(k) &= y(k-2) + y(k-1) + u(k)
\end{aligned}
\tag{3.27}
$$

The sequence for MPO prediction should be computed as follows:

$$
\begin{aligned}
\hat{y}(1) &= NA \\
\hat{y}(2) &= NA \\
\hat{y}(3) &= y(1) + y(2) + u(2) \\
\hat{y}(4) &= y(2) + \hat{y}(3) + u(3) \\
\hat{y}(5) &= \hat{y}(3) + \hat{y}(4) + u(4) \\
&... \\
\hat{y}(k) &= \hat{y}(k-2) + \hat{y}(k-1) + u(k)
\end{aligned}
\tag{3.28}
$$

From the previous examples, it can be observed that the first value that can be estimated is $k+1$, k being the maximum lag on the model. It re-

sults evident that for OSA predictions, the previous output measurement accounts for most of the prediction value, it is therefore easy to get good OSA predictions even with a poor bias model.

In the case of MPO predictions, the model is initialised using measurements from the dataset, once the maximum model lag has been reached by the estimations, all future predictions are based on previous model estimates. From the previous example, from the fifth prediction onwards, all model estimates are based on previous model estimates. If the model is not appropriate, the prediction error will accumulate as the prediction horizon increases. It is therefore necessary to always assess the model's performance using MPO and not OSA predictions.

A set of different models was estimated using the NARMAX methodology presented on the previous sections. From the 69 data points on the dataset, 57 were used for training the while the last 12 were used for validation. This partition comes from the fact that most oil companies plan their operations and forecasts one year in advance, therefore it is convenient to use the last 12 data points for validation. By doing so there is a partition of approximately 80% for training and 20% for validation, which is normal in many machine learning applications [53].

The models were estimated by increasing the number of maximum lags as well as the degree of the polynomial. From the estimated set of models, the MPO prediction performance was the main concern, based on the MPO prediction error the best models were further analysed.

NARX Model Validation

The performance tests that were used to evaluate the quality of the Mean Square Error (MSE) and the percentage of fit. The mean square error index is a good measure to evaluate the model's performance, it indicates the absolute difference between the model's forecasts and the measured data.

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 \tag{3.29}$$

Where n is the number of samples, and e is the difference between the corresponding forecast and measured data point.

$$e(k) = y(k) - \hat{y}(k) \tag{3.30}$$

A good model should produce a MSE as small as possible, indicating that the model's forecasts are close to the measured data.

The percentage of fit is a widely accepted index for evaluating the quality of a model's estimates, it is also known as Mean Absolute Percentage Error (MAPE). It is an intuitive index and is not scale dependent. The closer to 100, the better the quality of the model's forecasts, it is defined as [50].

$$\%Fit = 100\frac{1}{n}\sum_{i}^{n}(1 - \frac{|y_i - \hat{y}_i|}{|y_i|}) \tag{3.31}$$

Where n represents the corresponding sample y_n and prediction $\hat{y}_n$. This metric measures the model's error in terms of percentage. For example it is a lot more intuitive to say "the model is accurate 89% of the time" than "the model has an error of 14.563" if the interpreter is unfamiliar with the data. The usage of MAPE is very popular in forecasting applications like: weather

prediction, operations research, economy and applied statistics. A good review of the usage of MAPE is given by Neiting [73]. The performance and parameters of the estimated models are summarised on the following table.

Model	Polynomial Degree	Lags	# Of Inputs	Candidate Terms	Selected Terms	OSA Fit	MPO Fit	MSE OSA	MSE MPO
1	2	2	20	946	15	0.8915	0.6326	0.0166	0.1897
2	2	4	20	3655	15	1.0613	1.0796	0.0037	0.0063
3	2	6	20	8128	15	0.9968	0.9866	0.0000	0.0002
4	2	8	20	14365	15	1.2493	1.8164	0.0533	0.5714
5	2	10	20	22366	15	4.9007	24.2095	4.0728	144.1905
6	2	12	20	32131	15	1.1072	1.6822	0.0022	0.0903
7	2	14	20	43660	15	1.1007	1.5868	0.0015	0.0525
8	2	16	20	56953	15	1.0993	1.5829	0.0015	0.0524
9	2	18	20	72010	15	1.1606	2.0357	0.0041	0.1706
10	2	20	20	88831	15	0.8994	0.5101	0.0016	0.0389
11	2	22	20	107416	15	0.9249	0.6150	0.0009	0.0249
12	2	24	20	127765	15	1.2483	2.5071	0.0089	0.3267
13	3	2	20	14190	15	0.4826	-1.3202	0.3763	7.5659
14	3	4	20	105995	15	0.8922	0.7100	0.0115	0.0835
15	3	6	20	349504	15	1.1058	1.3484	0.0096	0.1041
16	3	8	20	818805	15	1.0266	1.1604	0.0003	0.0123
17	3	10	20	1587986	15	1.0315	1.1369	0.0003	0.0050
18	3	12	20	2731135	9	1.0747	1.5985	0.0011	0.0695
19	3	14	20	Too Many Terms	NA	NA	NA	NA	NA

Table 3.1: Polynomial NARX Models

For the present case study, the term selection algorithm was stopped at 15 terms or 95% of the SERR value. It was found that for most cases after the 15th term, the contribution of the following ones was so little that it would require more than 30 terms to reach 95% of the output variance, this is an indication that the subsequent terms might be noise terms that do not represent the dynamics of the system. Selecting a model with too many terms was found to lead to unstable forecasts. As a result, most 15-term models reached a SERR value close to 91%.

On Table 3.1 it can be observed how the number of lags and the maximum order of the polynomial function drastically increase the number of candidate terms.

According to the results, the best candidate models are models 2 and 3, since their performance indices for MPO predictions are the best. Based on these model's terms, there is an indication that it takes less than 6 months for the injected water to push oil into the surface, therefore models with further lags have poor performance. Some larger lag terms which are selected by the FROLS algorithm have high correlation with the output but incorrectly represent the field's dynamics. This is where physical interpretation of the system plays an important role. If incorrectly performed, the selection of inappropriate parameters will follow. This analysis is consistent with the field's dimensions since Scott is a relatively small field, and it's unlikely that it would take more than 6 months for injection to affect production [74].

It was found that third order models were really good for fitting the training data, but had bad generalisation results when performing on unseen

data as it can be observed on the following plot.

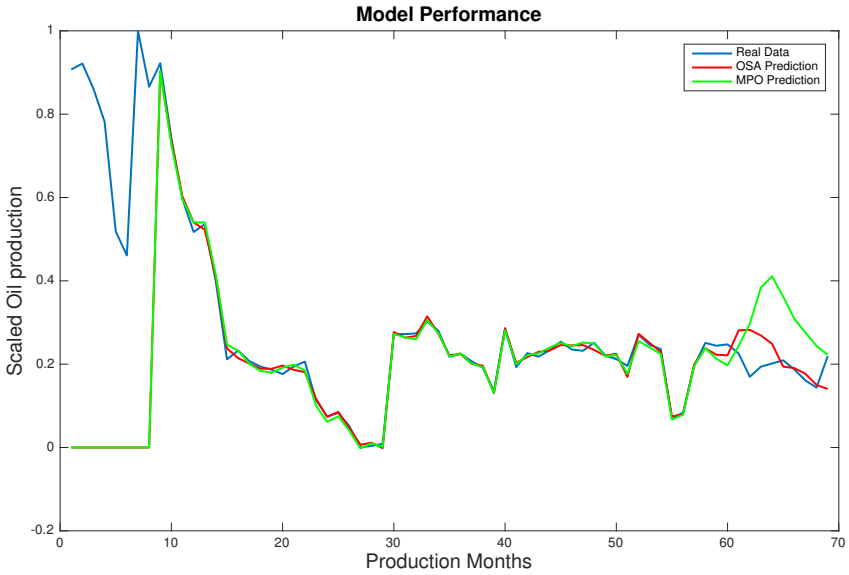

Figure 3.2: Third order Polynomial-Model 16 Performance

It results clear that after the 57th data point the forecast data quickly diverges from the measurements. Third order models are very likely to be over-fitted and should not be used for forecasting unseen scenarios on this dataset.

To determine which model is the most appropriate for describing the field's dynamics as well as forecasting a future scenario a quantitative analysis based on the performance indices is not sufficient. The following plots show how well models 2&3 can track the measured data.

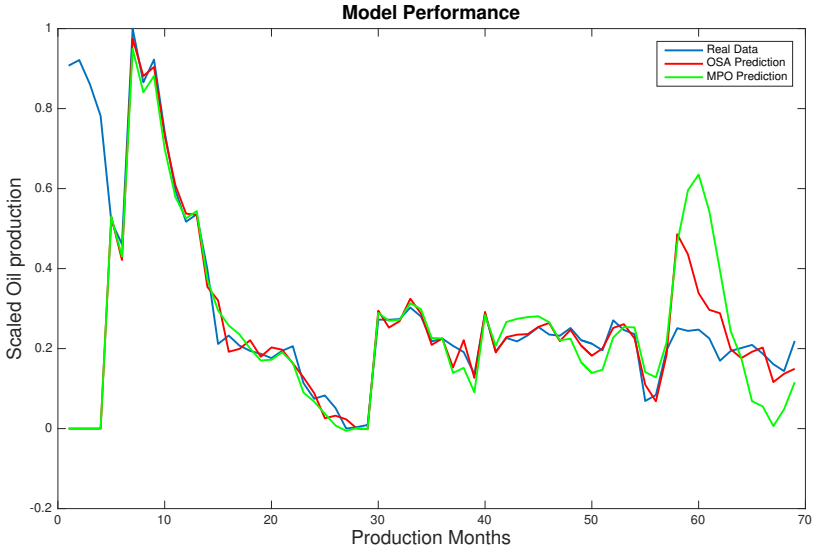

Figure 3.3: Model 2 Performance

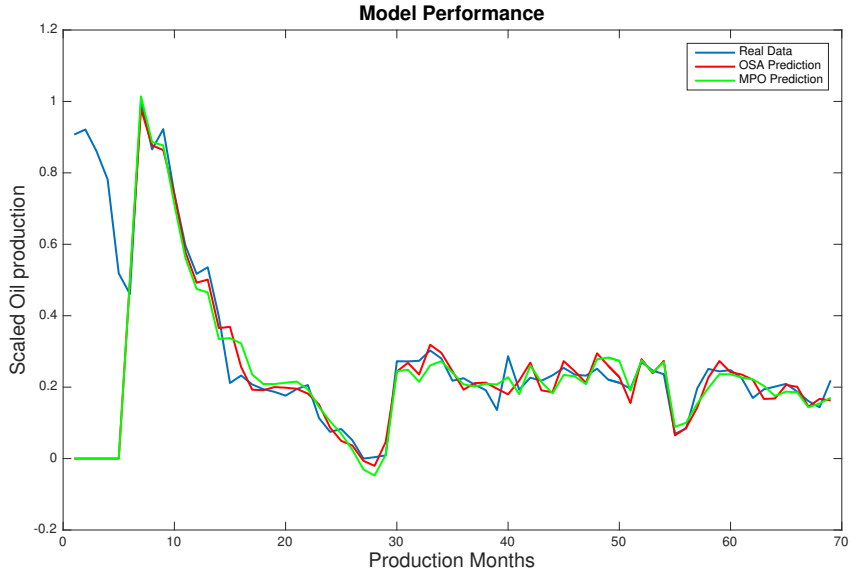

Figure 3.4: Model 3 Performance

It results evident that model 3 can track the measurements very well for both training and validation data. To fully validate model 3, the model's residuals were analysed. According to [50] a good model should produce normally distributed residuals which must lie within 95% confidence limits.

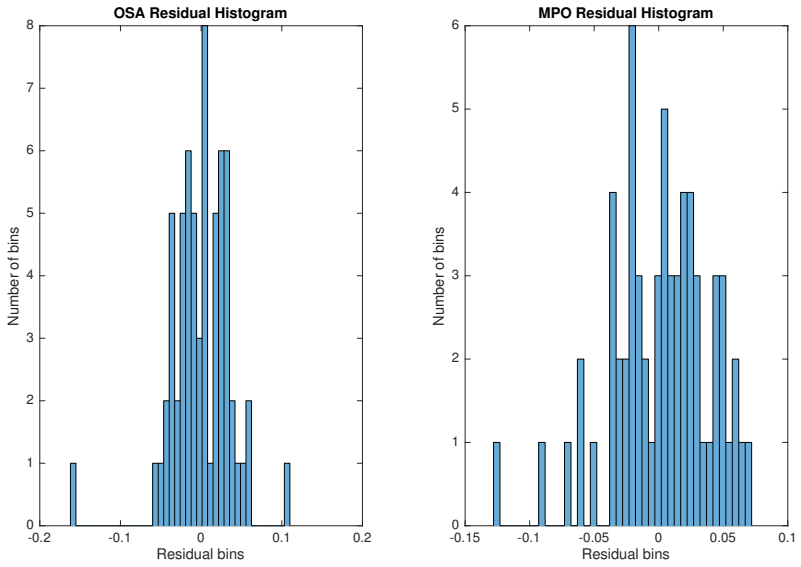

Figure 3.5: Model 3 Residuals Histogram

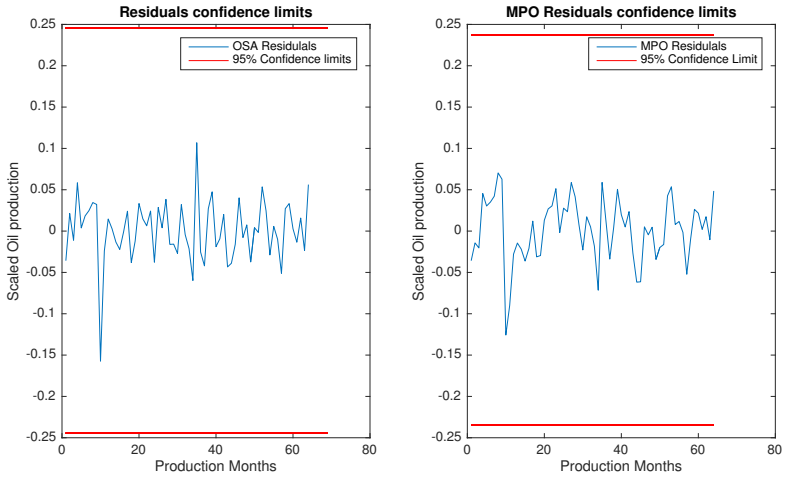

Figure 3.6: Model 3 Residuals Confidence Limits

As it can be seen on the previous diagrams, model 3 produces appropriate residuals and should be accepted for further analysis.

The equation for model 3 is:

Theta	Term	ERR
0.9800	'y(k-1)'	0.8480
1.1431	'u(k-4,2)*u(k-1,3)'	0.0205
-0.6783	'u(k-5,1)*y(k-1)'	0.0155
-0.2423	'u(k-1,3)*u(k-3,7)'	0.0059
0.3021	'u(k-5,3)*u(k-2,15)'	0.0043
-0.4351	'u(k-1,13)*u(k-4,14)'	0.0038
-0.7768	'u(k-1,18)*u(k-3,19)'	0.0024
0.3449	'u(k-5,1)*u(k-1,13)'	0.0020
-0.2140	'u(k-5,2)*u(k-1,3)'	0.0015
0.0766	'u(k-2,4)*u(k-5,16)'	0.0012
-0.4358	'u(k-1,3)*u(k-3,15)'	0.0012
0.2024	'u(k-2,2)*u(k-3,15)'	0.0012
-0.3660	'u(k-2,5)*y(k-5)'	0.0013
0.3352	'u(k-3,5)*y(k-5)'	0.0009
-0.1447	'u(k-3,2)*u(k-4,19)'	0.0007

Table 3.2: Polynomial NARX Model 3

The model terms should be interpreted as:

$$\theta * z(k - delay, var) \tag{3.32}$$

Where θ is a constant, and z is the corresponding output y or input u (given by var), $delay$ is the lagged sample the model has to use when producing a forecast. For example, the second term is:

$$1.1431 * u(k-4,2) * u(k-1,3) \tag{3.33}$$

The term must be read as: 1.1431 times injection well 2 with a time delay of 4 months multiplied by injection well 3 with a time delay of 1 month. The resultant model estimate is the sum of all term values.

From the model's structure it results clear which are the most significant inputs (injection wells). The terms are given in their order of contribution towards the output (oil production), the most important ones appear at the top of Table 3.2. In the case of the analysed production well, the frequency of appearance from the injection wells is given by the following table:

Injection Well	Frequency
1	2
2	4
3	5
4	1
5	2
6	0
7	1
8	0
9	0
10	0
11	0
12	0
13	2
14	1
15	3
16	1
17	0
18	1
19	2
20	0

Table 3.3: Injection Well Frequency-Model 3

From the model's structure it results evident that the most significant injection wells are wells 2,3 and 15. It is on these wells where changes on the injection rates would have a greater effect on production.

From the frequency table it can be observed that only 12 from the 20 injection wells appear in the model's equation. If the analysed production well was the only one in the field it would be advised to close the operation of those 8 injection wells which do not contribute towards production. In order to determine which wells should close operations, an individual model for every production well should be estimated. Based on the final model equation terms, those injection wells which contribution is small or none should be reviewed for future operations.

3.3 Conclusions

On the current chapter, it is shown how to extract meaningful information from a dataset containing production/injection rates coming from an operating oil field by using the NARMAX methodology.

The importance of data pre-processing is demonstrated. Without an adequate handling of outliers, missing values and data normalisation it is likely that the estimated models would lead to incorrect analysis and results. NARMAX models can provide a powerful tool for analysing a dataset. Specifically, polynomial NARX models offer flexible transparent equations which can be directly linked to the field's physical parameters. This enables the possibility of having a broader understating of the field's dynamics, which is not straightforward on traditional black box models. However, the selected parameters for estimated models must be chosen in accordance to the physical system. The resultant models must be validated using both quantitative and qualitative analysis.

The implementation of the FROLS algorithm provides very intuitive results since the model's terms are selected in their order of contribution, this has a direct relation to which injection wells are the ones with greater effect towards oil production. Based on these results, production engineers can take decisions which would lead to an efficient use of the resources, better planning for water injection and greater recovery rates.

Chapter 4

Novel Probabilistic NARX Neural Network Model Approach

In theory, a single layer neural network can approximate any continuous function [75]. However, in literature many authors prefer using several layers leading to better data fitting approximation at the cost of more complex structures. The structure selection process consists in determining an appropriate number of hidden layers, hidden neurons, selecting an appropriate activation function for the neuron perceptrons and chosing the correct weights between neurons. Once an activation function has been selected, the parameters within this function also have to be estimated, for example a Gaussian function has 3 parameters, mean, standard deviation and centre. As it has been mentioned, the selection of the the network structure and the number of parameters to estimate are non-trivial tasks that require considerable computational time. Some of the most popular training algorithms

can be found at [76] and [41].

4.1 Single Layer Multi-Scale Radial Basis Function Models

Obtaining a good reliable neural network model requires the estimation of a large number of parameters, consumes considerable time and the resulting model is a black box, where no direct interpretation of the physical parameters can be observed.

Motivated by the complexity of the problem, Billings and colleagues developed a simple single layer neural network model, this model structure is linear in the parameters and can be trained using conventional non-linear optimisation methods, or the well known FROLS algorithm [77] [78].

If a Gaussian function is used as the activation function, the case of a single-input single-output model with multiple centres, the function can be represented by the following expression [55]:

$$f(x) = w_1(e^{-\frac{1}{2\sigma}(x-c_1)^2}) + w_2(e^{-\frac{1}{2\sigma}(x-c_2)^2}) + ... + w_N(e^{-\frac{1}{2\sigma}(x-c_N)^2}) \quad (4.1)$$

For the multivariate case the previous expression can be extended as:

$$x_1 = \begin{bmatrix} x_1(1) \\ x_1(2) \\ ... \\ x_1(N) \end{bmatrix}, x_2 = \begin{bmatrix} x_2(1) \\ x_2(2) \\ ... \\ x_2(N) \end{bmatrix}, x_N = \begin{bmatrix} x_n(1) \\ x_n(2) \\ ... \\ x_n(N) \end{bmatrix}, y = \begin{bmatrix} y_1(1) \\ y_1(2) \\ ... \\ y_1(N) \end{bmatrix} \quad (4.2)$$

Arranging the terms on Equation 4.2, the model can be summarised as:

$$y(k) = \sum_{j=1}^{N} \varphi(x(k); \sigma, x(j)) + e(k) \tag{4.3}$$

The number of candidate centres is given by $j = N$, the term $e(k)$ represents the error between the data and the predictions generated by the model. The weights can be estimated by solving a simultaneous system of linear equations.

$$\begin{bmatrix} y_1(1) \\ y_1(2) \\ ... \\ y_1(N) \end{bmatrix} = \begin{bmatrix} \varphi_1 1 & \varphi_1 2 & ... & \varphi_1 N \\ \varphi_2 1 & \varphi_2 2 & ... & \varphi_2 N \\ \vdots & \vdots & ... & \vdots \\ \varphi_N 1 & \varphi_N 2 & ... & \varphi_N N \end{bmatrix} \begin{bmatrix} w_1 \\ w_w \\ ... \\ w_N \end{bmatrix} + \begin{bmatrix} e(1) \\ e(2) \\ ... \\ e(N) \end{bmatrix} \tag{4.4}$$

Several studies show radial basis functions with a single scale (standard deviation) have limited performance for dynamical modelling, a detailed explanation about this issue can be found on [79], [56], [80]. This limitation is because not all input variables contribute to the same degree to the system's output.

If large-scale values are selected, global dynamics are captured. On the other hand selecting small-scale values only captures local data behaviour. In literature, this problem has been solved by adaptively adjusting scale values, but this approach leads to complicated practical implementation and optimising the scales size is computationally expensive [81] [54].

The challenge has been solved by using different scales and then estimating a set of candidate terms. The most significant terms and corresponding

scales can then be selected using the FROLS algorithm.

A multi-input, multi-scale single output Radial Basis function is described by the following expression [55]:

$$\hat{f}(x(k)) = \sum_{i=0}^{I}\sum_{j=0}^{J}\sum_{m=1}^{N_c}\theta_{i,j,m}\phi_{i,j,m}(x(k);\sigma_m^{i,j},c_m) \tag{4.5}$$

If a Gaussian function is selected as an activation function the previous expression changes to:

$$\phi_{i,j,m}(x(k);\sigma_m^{i,j},c_m) = exp\left[-\sum_{r=1}^{n}\left(\frac{x_r(k)-c_{m,r}}{\sigma_{m,r}^{(i,j)}}\right)^2\right] \tag{4.6}$$

The 1/2 term on Equation 4.1 is now irrelevant since many scale values will be evaluated, multiplying the standard deviation by a constant does not change anything.

As it can be observed on Equation 4.6, the number of candidate terms is dependent on the number of centres and the number of candidate scales that will be evaluated for every centre's value. Therefore, the selection of the function's centres is another critical parameter. When using a time series approach, lagged variables must be considered, the previous equation can be written as:

$$\begin{aligned}\phi(x(k);\sigma,c_j) = exp\big[&\left(\frac{y(k-1)-c_{j1}}{\sigma}\right)^2 + ... + \left(\frac{y(k-ny)-c_{j,ny}}{\sigma}\right)^2 \\ &+ \left(\frac{u_1(k-1)-c_{j1}}{\sigma}\right)^2 + ... + \left(\frac{u_n(k-n_{nun})-c_{j1}}{\sigma}\right)^2\big]\end{aligned} \tag{4.7}$$

The number of candidate scales and their size is something that must be selected. For normalised data (from 0 to 1) a good starting scale value is $\sqrt{10}$ [82]. This value has proven to be large enough to capture global dynamics while local dynamics can be captured by smaller scales.

The number of initial candidate terms is:

$$N_s = (I+1)(J+1)N_c \tag{4.8}$$

Where $I+1$ is the number of selected output scales, $J+1$ is the number of input scales, N_c is the number of selected centres.

The FROLS algorithm can be used to train the model [56]. The methodology provides a trade-off between simple single layer radial basis functions and complex network structures where non-trivial algorithms have to be implemented in order to estimate the model's parameters.

The methodology can be summarised on the following steps:

1. Select the candidate kernel centres, either all points in the dataset or a few by a cluster algorithm.
2. Heuristically assign the number of scales and their values.
3. Use selected centres to form a 2D grid.
4. Form a MSRBF where the inner parameters are defined by the 2D grid.
5. Convert the MSRBF network into a linear in the parameters form.
6. Select the most significant terms using the FROLS algorithm.

The selection of the number of centres can drastically increase the number of candidate terms. If the size of the dataset we are working with is not excessive (Less than 10,000) we can then place the radial basis centres at each of the points from our dataset.

4.2 Implementing MSRBF Models for EOR Modelling

Multi-scale radial basis function NARX models have shown their capability on diverse applications such as: liquid level through a DC motor, artificial complex systems and heat exchanger system [83], [84]. However, the model's benefits have not been explored in the area of history matching. The challenge of our application is that the oil field under analysis has 20 injection wells, so we have to consider at least 20 inputs, this means that we have to select the centres for at least 21 functions (20 inputs + 1 output) compared to the Single-Input Single-Output (SISO) systems where the model has been used.

Since the analysed data set is not very large, all the data points were used as candidate centres (57 points for training the model). Future production values depend on the previous state of the reservoir. Therefore, a time series approach must be applied. The model's structure that was implemented is the one shown on Equation 4.7.

As a rule of thumb following a suggestion of Chen, all data values were normalised from 0 to 1, a maximum value of $\sqrt{10}$ was used for all the candidate scales, further smaller values were estimated as follows [82]:

$$\sigma_{y,u} = 2^{-i,j}\sigma_{y0,u0} \tag{4.9}$$

Where σ_{y0} and σ_{u0} are the maximum output and input scale values, respectively. The number of candidate scales is given by i for the output scales and j for the input ones. During the construction of candidate terms, a set of different combinations between the input and output scale values are tested as on Equation 4.8.

There is no rule for determining what are the appropriate scale values or number of lags, these parameters have to be determined heuristically.

A set of different models was estimated to determine the best parameters. The number of candidate scales was set to 5 ranging from $\sqrt{10}$ to $\sqrt{10} * 2^{-4}$. The number of lags was increased heuristically until no improvements on the model's predictions were observed. The following plot shows the how the model's performance changes according to the number of output lags.

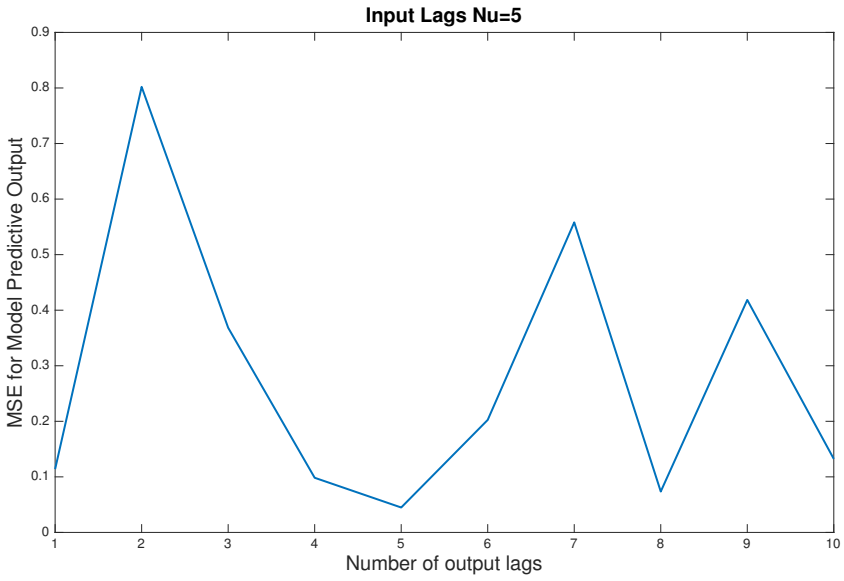

Figure 4.1: Output Lag Estimation

As it can be observed on the previous plot, the model's performance seems to change randomly and independently of the number of output lags. However, according to the plot, the model containing 5 output lags outperformed all others.

After selecting $ny = 5$ as the reference staring point, the same procedure to estimate an appropriate number of input lags was followed. The number of input lags was gradually increased to 10, from which the optimal value was observed.

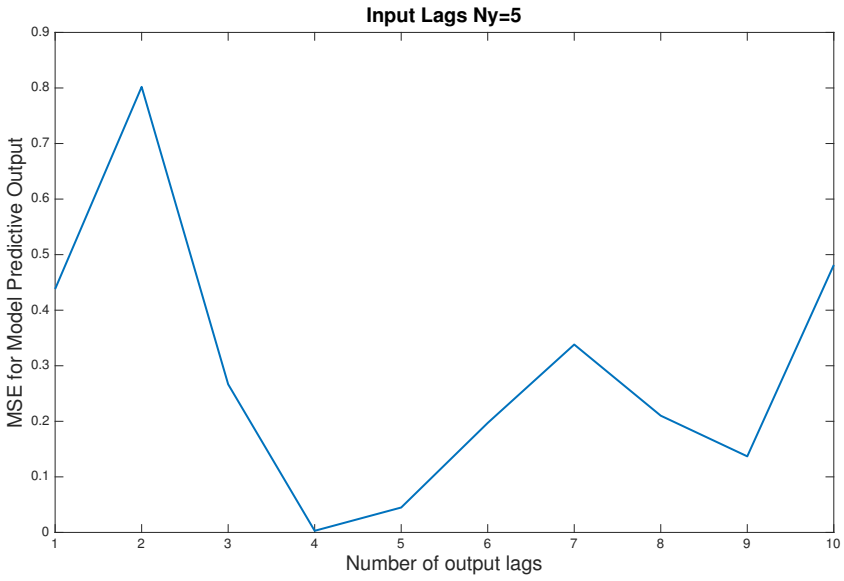

Figure 4.2: Input Lag Estimation

From Figure 4.2, the optimal input number of lags is 4 as it minimises the mean square error. According to the performed tests, the best model should be the one with $ny = 5$ and $nu = 4$. This model's performance is shown on the following plot.

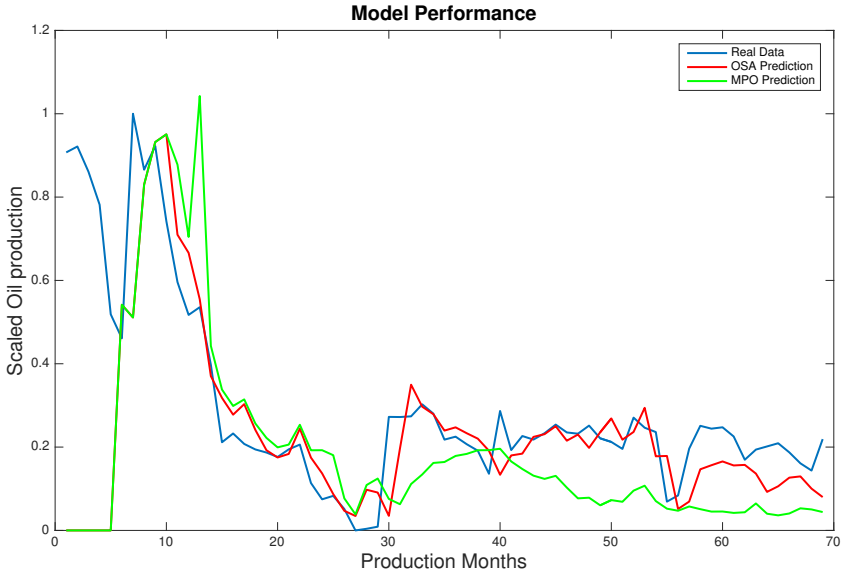

Figure 4.3: MSRBF Model ny=5 nu=4

As it can be observed on Figure 4.3, the model's performance is quite poor. One step ahead predictions do follow the measurements with an offset while the model predictive output is far from the measured dataset. See Appendix A for model's structure.

After analysing the structure from the generated models an interesting pattern was found. The FROLS algorithm selected those terms with large input scale values, this happens because on a time series model the most significant term is $y(k-1)$, as a consequence one step ahead predictions can be very good if most of the model's weight is placed on this term. On the other hand, by doing so the model predictive output becomes unstable leading to poor long term prediction results.

Selecting input terms with large input scale values means the inputs don't affect the output, in our case it would mean that injection rates do not affect oil production which is an incorrect analysis. This is unacceptable; the model must be able to represent the system's dynamics in the long term, and account for the system's inputs.

There are two options for increasing the input's effect on the model's output, one is reducing input scale values and the other is by increasing the number of lags on the input variables.

As it has previously been presented, determining an appropriate number of lags is a non-trivial task for non-linear systems. For linear systems, cross correlation or partial cross correlation functions can help to determine an appropriate number of lags [33].

For the single-input single-output case, tuning the model is a straightforward process since it is relatively easy to test many scales with a set of input-output lag combination. In the case of EOR the story is a bit more complicated. The number of input variables can can be quite large (one for every injection well), therefore the number of possible combinations is huge. The number of possible combinations is given by the following equation:

$$N = [(\sigma_{u1})(n_{u1})][(\sigma_{u2})(n_{u2})]...[(\sigma_{un})(n_{un})][(\sigma_y)(n_y)] \tag{4.10}$$

Where σ are the corresponding scales and n the number of lags. For example, using 6 delays for all input variables and 2 delays for the output and 5 candidate scales leads to the following number of combinations:

$$N = (5 * 6)^{20} * (5 * 2) = 3.4868(10)^{30} \tag{4.11}$$

Testing all the possible combinations is almost impossible, there are simply too many for a computer to process. Before analysing the dataset we can't know what the most appropriate model structure is and it'll be very easy to end up with a poor model.

One of the main issues MSRBF models have to deal with is how to manually adjust the dependency on certain input variables if we do not know anything about the physical system properties. In many cases we may actually want to understand how the variables are related to each other by estimating a model, by using a MSRBF model this task becomes a very computationally expensive task if the number of inputs is significantly large.

4.3 Novel Pruning Method for NARX MSRBF Models

In literature, most of the applications where single layer MSRBF NARX models are used deal with a low number inputs. It therefore results relatively easy to estimate a decent number of candidate terms which will lead to a good model. If the number of inputs is relatively low, the presented methodology on section 4.2 can directly be implemented [54],[85], [78].

If the analysed system has multiple inputs, the methodology becomes either computationally expensive or inefficient leading to poor models if not enough parameters are selected. In such cases the methodology can not be

used off the shelf.

Motivated by the simplicity of MSRBF NARX models, a new variant was developed to produce appropriate forecasts when dealing with a high number of inputs.

In machine learning literature, it has been proved that model combination in most cases improves the performance of predictions. The idea is simple, combine the predictions of many models and the final prediction in average will be better than the ones from a single model.

In theory, models with different architectures would compensate for the errors of other models. In practice, this approach is too complex since tuning the parameters of all the models is nearly impossible [86].

On a traditional multi-layer neural network structure, complex input-output relations can be built by estimating a set of weights between the layers. This process can be time consuming and computationally expensive. If the dataset is not very large, it is very likely that the network will be affected by noise, therefore over-fitting may happen.

Finding an appropriate parsimonious network structure can be a challenging process. If the network is too large, the estimation of unnecessary parameters will happen, if the network is too small its ability to create estimates will be compromised. In literature, it has been found that small networks are the ones with the best generalisation performance but tuning the parameters can be a complex task. To deal with the mentioned issues

when working with neural networks, researchers have developed diverse techniques such as pruning the original structure, a good summary of the most popular methods can be found on [87]. Regardless of the pruning algorithm, the concept of combining of estimates from simple model structures in all reported cases produced better generalisation estimates than a single model containing all the inputs and hidden layers.

The idea of combining models comes from natural evolution. In nature, reproduction involves taking half of the genes from each parent producing a random mutation. This approach has proven to be the most effective for evolution, where the weaknesses from one parent are compensated by the strengths of the other. On the other hand, asexual reproduction replicates a child from a parent coping all its properties, this process is faster than combinational reproduction.

A direct analogy can be made, the more combinations the better and more resistant to diseases the children would be. Therefore, the more models that are estimated, the better and less sensitive to noise the output would be. In the case of asexual mutation, the analogy would be to over-fit a model, which will only replicate the training data but can be estimated in little time.

In order to deal with the problem of mixing and estimating models with different structures, Srivastava and colleagues presented a simple solution for multiple layer neural networks. The methodology is known as "Dropout", where some of the links and percetrons on the network are randomly deactivated [86].

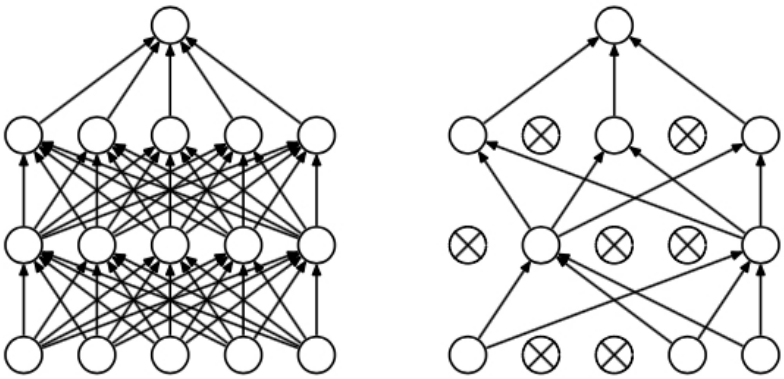

Figure 4.4: Drop-out Diagram

[86]

By eliminating elements on the network, less parameters have to be estimated leading to less complex structures, less parameters to estimate requiring less computational time when building single models.

The idea is to build many single drop out models and then use an ensemble structure to forecast with all of them, combine the forecasts and produce a final prediction.

For the case of MSRBF NARX models there is only one layer, therefore only inputs can be deactivated from the structure. The dropout diagram for these models can be appreciated on the following figure.

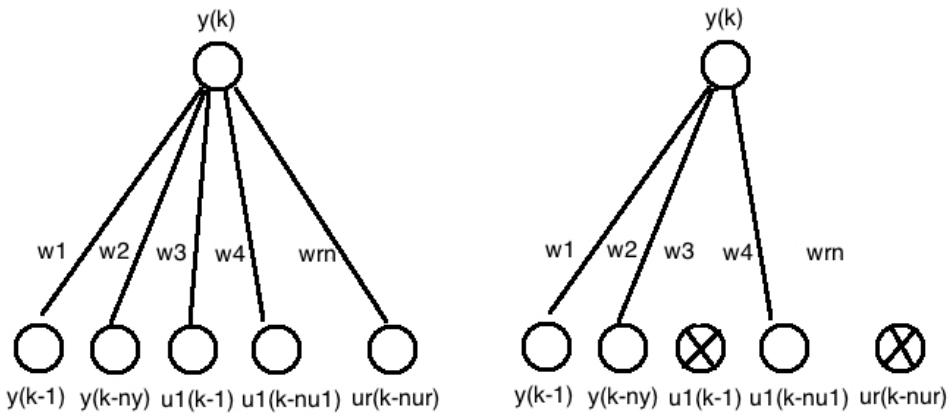

Figure 4.5: MSRBF Dropout Diagram

For the analysed case study, there are at least 20 elements on the network, representing every lagged input. The idea is to estimate different models by randomly changing which inputs should be considered on the network structure.

Srivastava, suggests using a probabilistic method for selecting which inputs should be present when estimating a new model [86]. There is not an "optimal" value for this probability, in literature it is suggested to use a value between 0.5 and 1. But this value is based on specific experiments and model structure.

Since this probabilistic approach has never been used for MSRBF NARX models, the suggestion is vague. By using a value higher than 0.5 we would expect to have most of the inputs active. This approach would help to reduce the number of candidate terms, but it would still be very high.

When estimating models for history matching, it is very likely that the analysed field has more than 10 injection wells. So depending on the number of inputs, the "activation" probability value can be adjusted. A problem of using a low number of inputs is that the estimated models might not be able to replicate the system's dynamic and poor forecast estimates will be produced. To deal with this problem, a value of 0.5 was used on the present case study, by doing so the number of candidate terms to be estimated is substantially reduced.

Instead of changing the scales as on Equation 4.7, a number of 3 random scales values between 0 and $\sqrt{10}$ were used when generating candidate terms.

As it has previously been mentioned, estimating an appropriate number of lags for non-linear systems is a non-straight forward process. By using the results from the polynomial model, the maximum lag values set as the ones from the best polynomial and MSRBF models.

The new proposed methodology can be summarised as follows:

1. Define maximum number of lags (By correlation analysis, heuristically or by any other method).
2. Randomly select the number of scales and their values.
3. Use a probability function of 0.5 to activate approximately half of the inputs (This value can be tuned depending on the dataset).
4. Use selected centres to form a 2D grid.

5. Form a MSRBF where the inner parameters are defined by the 2D grid.
6. Convert the MSRBF network into a linear in the parameters form.
7. Select the most significant terms using the FROLS algorithm.
8. Compute model forecasts. If model passes the requirements, keep it, if not, discard it.
9. Combine forecasts from all successful models.
10. Repeat steps 1-9 until no further improvement is observed,

The requirements from element 8 on the list can include: MSE, % of fit, residuals distribution or another performance criteria.

The number of required models depends on the analysed dataset and must be determined every time a new system is analysed. For the Scott Field case study, the requirements were: a MSE value of 0.001 for the whole dataset and a maximum gain of 10% for the validation data, both for the one step ahead prediction and the model predictive output. These parameters can be adjusted depending on the required accuracy.

The following plots show how the MSE error changes according to the number of models on the ensemble structure.

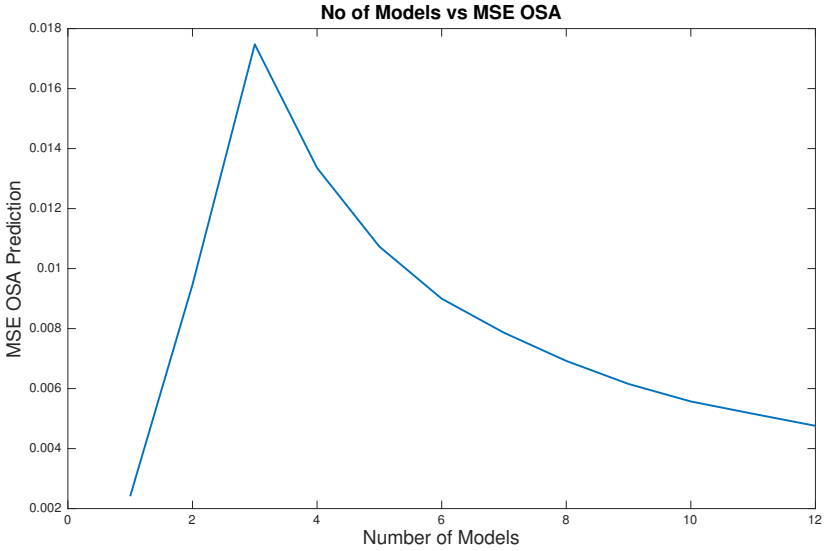

Figure 4.6: No of Models vs MSE OSA

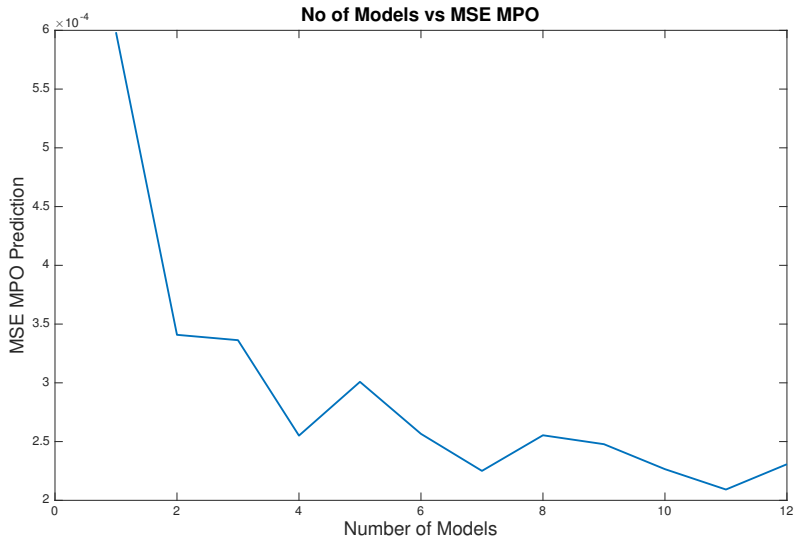

Figure 4.7: No of Models vs MSE MPO

As it can be observed on the previous plots, the mean square error reduces as the number of models on the ensemble increases. For the Scott Field example, it appears that the MSE converges around 0.004 for one step ahead predictions and 0.0025 for the model predictive output estimates. According to the results, after combining 12 models the MSE reaches a steady value.

After running several tests, it was found that by using the selected parameters, only 0.6% of all the generated models passed the requirements. Since it is known how long it takes to create a new model it is therefore possible to know how long it will take to create an appropriate ensemble. For the present case study, it takes approximately 40 seconds to estimate a model. So about 1.85 hours to find one that complies with the requirements. Therefore, in order to come up with a decent prediction it may take one day of computing resources.

The produced forecasts from each of the 12 models are displayed on the following plots.

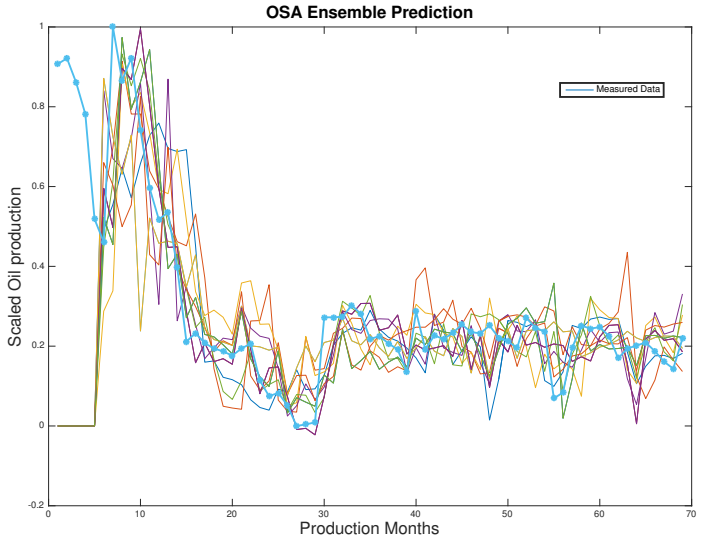

Figure 4.8: Ensemble OSA Predictions

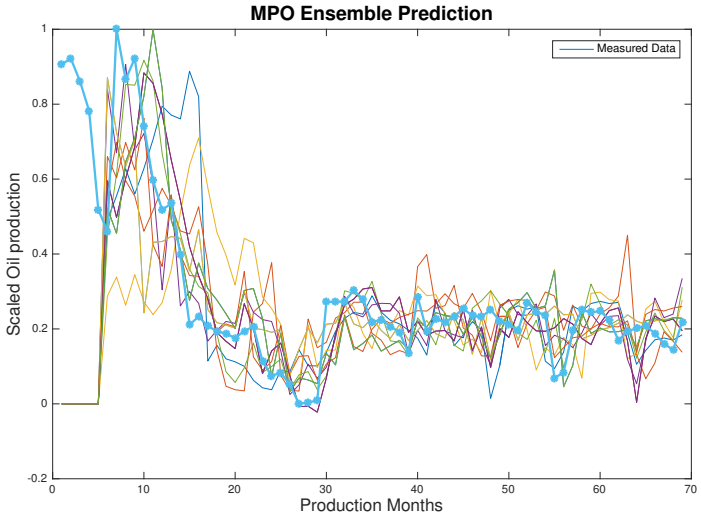

Figure 4.9: Ensemble MPO Predictions

As it can be observed on the figures 4.8 and 4.9, all model predictions follow the trend of the measured data. But some models can track it better depending on the operating point. The final ensemble prediction is shown on the following plot.

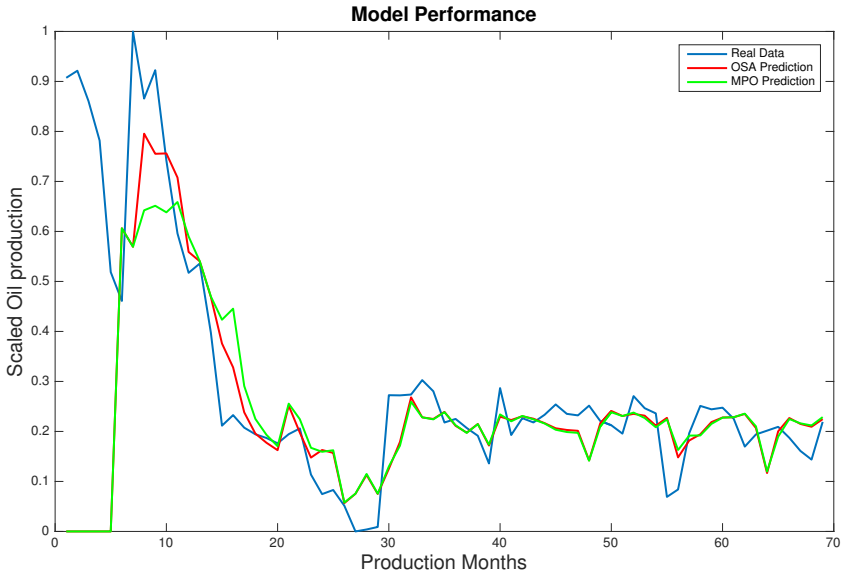

Figure 4.10: Ensemble Prediction Performance

From figure 4.10 it results evident that the prediction performance highly improves by using the suggested dropout methodology compared to the results shown on Figure 4.4. Quantitatively, the prediction performance for the validation data is summarised on the following table:

Model	MSE OSA	MSE MPO
Single 20 Input Model	5.38E-03	1.44E-01
12 Drop Out Model	9.17E-05	4.67E-04

Table 4.1: MSRBF Model Performance Comparison

As it can be observed on the previous table, the estimates' performance increased by approximately 59 times for one step ahead predictions and 300 times for the iterative model predictive output. The performance could increase even more if more models are built and considered on the ensemble. However, there is a trade-off, the more models that incorporated the less significant the improvement is, therefore a good approach would be plotting the MSE or any other performance index compared to the number of models on the built ensemble.

The following table shows the MPO prediction performance for the different type of models estimated on this thesis.

	Polynomial NARX	MSRBF NARX	Multi-Layer NN
MSE MPO	2.32E-04	4.68E-04	1.25E-03

Table 4.2: Model Performance Comparission

The prediction performance obtained by the polynomial and pruned MSRBF NARX models is higher than the one obtained by the standard multi-layer neural network as it can be seen on the following table.

4.4 Conclusions

Single layer multi-scale radial basis functions have proven to be very effective for finding a relation between input and output time series variables. The main advantage of this approach is the use of a simple training algorithm like the FROLS algorithm for selecting the most significant terms instead of complex non-linear optimisation methods.

The original multi-scale single-layer model by Chen and colleagues is easy to implement when the number of inputs is not very large [67]. However, when dealing with a high number of inputs like EOR modelling, the methodology requires the estimation of a huge number of candidate terms, demanding large computational resources. At the end, the methodology might end up with a poor performance model if the chosen parameters are not appropriate.

By randomly deactivating inputs and training models with simpler structures, the number of estimated parameters for a single model drastically reduces. This results in accurate forecasts even when structure of the models is not optimal.

Chapter 5

Estimating Future Uncertainty

Uncertainty is a concept that has to do with the degree of being sure what will happen in the future. History matching shares the same objective by adjusting models which enable us to look into future production scenarios. Based on those scenarios, take decisions that will increase oil production performance [88].

The only certain thing about forecasting future states, is that uncertainty will happen [89]. Uncertainty tells us how far we may be from the real value. Uncertainty might be random (having no pattern), fuzzy (variable level given by descriptions and not numeric values) or incomplete (coming from incomplete records) [90].

5.1 Feature Selection Comparative Study

Feature selection has the objective of selecting a subset of variables, discarding those variables that are not relevant to the problem understudy. For

EOR modelling, determining which injection wells are the most significant is a matter of great importance. Having active injection wells that do not contribute towards oil production has a significant cost and negative impact.

A variant of feature selection known as variable ranking has the objective of rating input variables according to their significance in the problem understudy. The results obtained by these methods are similar to the ones obtained by the final polynomial NARX model equation by using the FROLS algorithm.

Considering a set of n input variables $x_{k,i}(i = 1, .., n)$ and one output variable y_k, variable ranking uses a score function $S(i)$ to sort the input variables in a decreasing order according to their significance [91]. In literature, diverse methodologies for selecting the most significant variables are available, however some use what is known as feature transformation. These methods, are not recommended for problems where the meaning of the features is important as in EOR modelling. Therefore, these methods were not analysed.

The list of available methods without transforming the existing features into new features includes: All possible subsets selection, stepwise regression, bagged decision trees and regularisation [92].

All possible subsets selection works by generating models using all the possible input combinations and then estimating a model for each combination. This brutal force method works well for problems where the number of input variables is not very large. For the problem of EOR the number

of input variables can be significantly large, for example if a 6-month delay case wants to be analysed for 20 injection wells, the number of possible combinations rises to:

$$N = 2^{20*6} = 1.1529 * 10^{18} \tag{5.1}$$

This number of combinations is too high for a computer to handle, making the approach infeasible.

Decision trees work by estimating sets of models that include only a fraction of the input variables, the variables are ranked according to their effect when being present on each of the ensemble models. Stepwise regression works by sequentially adding or removing input variables until an acceptable model is estimated.

For the present case study, the selected method was sequential feature selection. This approach has the advantage of being able to use any type of model structure. This is very important since non-linear significance of input variables can be assessed. On the other hand, correlation based methods can only measure linear contribution of from the input variables, a good review of the most popular methods is given by Kudo [93].

Sequential feature selection can be forward, where features are added until no improvement is observed. Or backwards, where the methodology starts with a model containing all features, and features are sequentially removed until an acceptable set is found. For the present study, the forward strategy was chosen, since for EOR modelling it is more intuitive to find injection

wells from the most to the least significant.

The sequential feature algorithm can be summarised on the following steps:

1. Start with an empty set $Y_0 = \{0\}$
2. Select the next best feature $x = argmax[J(Y_k + x)]$
3. Update $Y_{k+1} = Y_k+, k = k + 1$
4. Go to step 2 until $C_k \leq Metric$

Where C_k is the value held by a criteria variable at iteration step k. Variable $Metric$ is a threshold value for determining when to stop the search of features x.

For the presented analysis, the objective is to compare the results obtained by the polynomial NARX model. From Tables 3.2 and 3.3, we can see that only 12 of the 20 injection wells were selected by the FROLS algorithm. For comparison purposes the sequential feature selection process was stopped at 12 features. The following flow diagram shows the criteria for selecting the most significant injection wells.

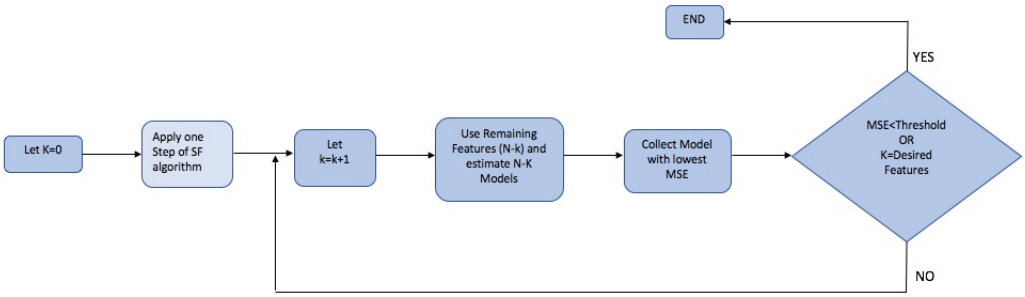

Figure 5.1: Flow Chart For Sequential Feature Selection Algorithm

The models shown on the previous diagram were estimated using 13 hidden neurons, one hidden layer and 12 delays for each variable. These parameters are the same that were used on section 2.5.1 to produce the best performing multi-layer neural network model.

Due to the differences that might exist with the initial conditions when training the multi-layer neural network using the Levenberg-Marquardt algorithm, a set of 10 different neural networks was estimated for every sequential feature addition. The mean of the 10 generated MSE values was used as the criteria for selecting the most appropriate feature to add into the model. To rank the contribution of the selected inputs, a statistical analysis was performed.

The feature selection process was run 20 times. This process is necessary for observing which features are selected first over a set of many experiments discarding the errors that could appear from running a single test. The number of runs was selected by observing changes on the order of selection from the parameters. The following plot shows the change in ranking when

selecting one of the features. As it can be observed the values converge after 12 experiments.

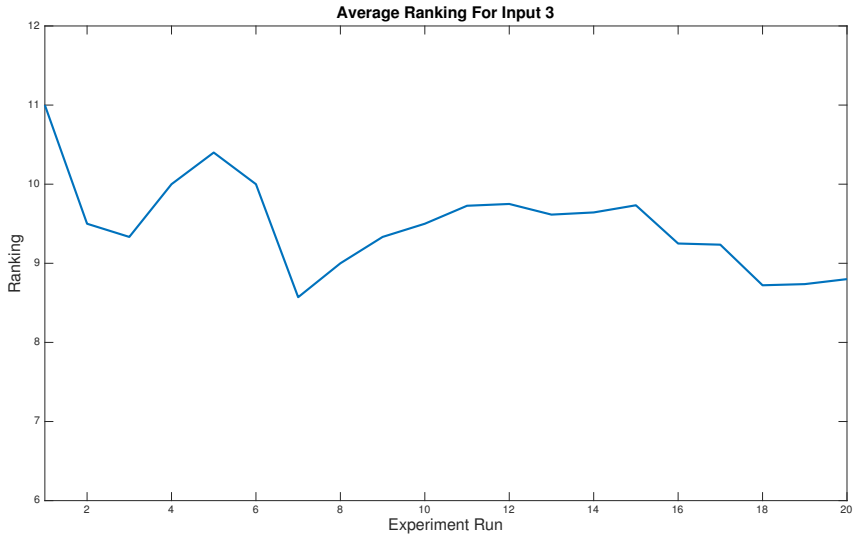

Figure 5.2: Moving Average Ranking For Input 3

The required number of models to be estimated using the previous approach can be quite large depending on the number of original and desired features to obtain. The following table shows the number of models to be estimated for every test.

Selected Features	Remaining Features	No of Estimated Models
1	20	200
2	19	190
3	18	180
4	17	170
5	16	160
6	15	150
7	14	140
8	13	130
9	12	120
10	11	110
11	10	100
12	9	90
Total		1740

Table 5.1: Number of Estimated Models

Since it was necessary to run the test 20 times, the total number of estimated models is 34800 (1740*20). This an inconvenience of the sequential feature selection method, it is very computationally expensive requiring long time to obtain the results.

The following table shows the computational time required to estimate all the polynomial terms and select the 15 most significant terms as well as the time required to run the analysis for sequential feature selection using 10 network estimations per feature and running the experiment 20 times.

	Polynomial	Multi-Layer NN 12 Features
Model Estimation	0.87 Sec	29893 Sec
Terms Selection	7.61 Sec	13303.8 Sec
Total	8.48 Sec	43196.8 Sec

Table 5.2: Feature Selection Time Comparison

The results shown on Table 5.2 were obtained running the analysis on Matlab 2015b using an Intel i5 1.3hz processor. The difference in processing time is huge, it takes 5094 times more to rank input variables using the sequential feature selection algorithm.

The significance of the input variables is presented on the following table. Ranking values are given in a descending order, 1 representing the most significant variable.

Input	Rank NN	Rank Poly	Difference
1	11	4	7
2	7	1	6
3	2	2	0
4	1	7	6
5	12	11	1
6	18	0	0
7	15	9	6
8	9	0	9
9	16	0	0
10	14	0	0
11	19	0	0
12	6	0	6
13	4	5	1
14	5	10	5
15	17	3	14
16	13	7	6
17	3	0	3
18	10	12	2
19	8	6	2
20	20	0	0

Table 5.3: Input Variable Ranking

The rank of the polynomial input variables was estimated by considering the number of times a term was selected by the FROLS algorithm as well as its relative position on the model. The weights were given a corresponding

factor of 14 depending on its position from Table 3.2. A value of 14 was chosen since there are 14 positions an input variable can take (14 terms including input variables). The weights can be computed as follows:

$$W(x_i) = L_1 + L_2 + ..L_k \tag{5.2}$$

Where X_i is the corresponding input variable i, and L_k represents the number of times the input variable appears on the model equation. For example the corresponding weight for input 1 is:

$$W(x_1) = 14 + 8 = 22 \tag{5.3}$$

Input 1 appears twice on the final polynomial model. First on the second position ($L_1 = 14$) and then on the eighth ($L_2 = 8$).

According to the results shown on Table 5.3. The models agree on the ranking of 6 from the 20 input variables. For most cases, there is a good match between the two approaches, existing small differences on 11 variables. The biggest difference in ranking is on input variable 15. This difference might be due to how the sequential feature algorithm determines which variable is the most significant according to the prediction error. After performing a full analysis by observing the final equation for all oil production well models, injection wells were ranked as follows:

Injection Well	Add Up	Add Up Ranking
1	67	1
2	44	13
3	28	17
4	61	3
5	61	3
6	20	18
7	59	5
8	32	16
9	55	7
10	39	14
11	54	9
12	53	10
13	55	7
14	53	10
15	37	15
16	46	12
17	56	6
18	11	19
19	63	2
20	0	20

Table 5.4: Injection Well Ranking

From the previous table it is clear that the most significant injection wells are wells 1,19,4 and 5. For full ranking results see Appendix B. The Add up column shows the weighted sum from an injection well appearance

as on Equation 5.3 for all production wells. The Add Up Ranking column shows the ranking for every injection well on the field based on the Add Up column, the ranking order goes from 1 to 20, 1 being the most significant.

The following plot shows how the prediction error changes according to the number of features that are integrated into the multi-layer neural network.

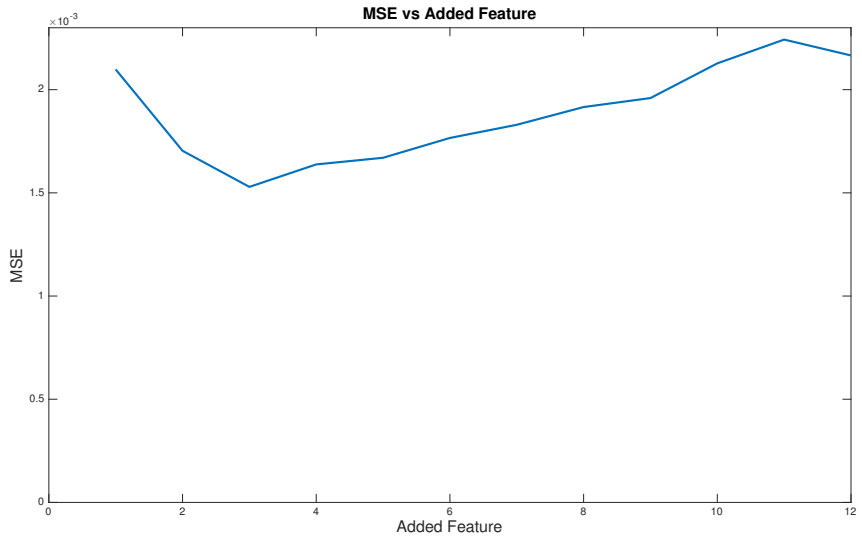

Figure 5.3: MSE vs Added Feature

From Figure 5.3 the MSE does not always decrease as new features are added into the model. If the order of variable selection changes, the subsequent models will also differ from other possible combinations which will result in a non-optimal solution. Since we don't know what the true answer for this problem is, we can only say that the presented methods should only be used as a diagnostic tool when ranking injection wells. Input variable

ranking is a problem understudy for industrial applications, no method has been capable of providing a deterministic result [50], [94].

5.2 Introduction to Risk Analysis

The concept of risk is very popular in modern finance. Every time there is a trading activity there is a certain level of risk [95]. Risk is defined as threat of a loss of time that can't be regained [96]. In other disciplines this concept is not very widespread, but this does not mean it is not present.

On this chapter, it is shown how risk analysis tools developed in mathematical finance can be applied to our NARX models and the problem of estimating future oil production by means of water injection. We are always exposed to external factors that can take us to a different scenario from the one we expected [97]. As long as we can't be truly sure of what will happen in the future, we will be exposed to a certain degree of risk. Forecasting a future scenario is possible since time series models can relate dependencies on the analysed variables from previous observations [98].

Uncertainty in predictions begins with the data we use for building our models, if we have bad data, we will have a poor model leading to bad predictions. There are several issues related to data collection that can happen on a real case study, this includes: inaccurate measurements or incomplete sets of observations. Both issues are susceptible to human error, faulty instruments and incorrect sampling period [98]. The difficult part is knowing how inaccurate or reliable the data is, we therefore need a mathematical description to measure this parameter. Using intervals provides a good description.

For example, if variable x can hold values between the interval $[x_1; x_2]$ we can say that the uncertainty of x lies within this range. How likely it is for x to be closer to x_1 or x_2 has to do with a membership function that describes the probability of the distribution. This is known as a fuzzy variable [98].

$$x = x, \nu_x(x) | \mathbf{x} \in \mathbf{X} \tag{5.4}$$

Uncertainty on the previous equation is represented by the membership function $\nu_x(x)$.

Now that the concepts of risk and uncertainty have been defined, we can make use of the techniques developed in risk analysis. As a definition risk analysis tools determine which factors can have a greater impact on the project we are analysing [99]. In our case these tools provide information about the input variables' effect on the output. There are two main divisions within risk analysis: qualitative and quantitative methods, which combined form a risk assessment.

5.2.1 Qualitative Methods

Qualitative methods are based on intuition, previous experience or simply a judgement for decision making. These methods are normally used when numerical data is not available or the problem's uncertainty is so low that there is no need for a mathematical approach.

The analysis qualitative methods provide is very important. At its first stage, qualitative methods identify what the risk sources could be. Risk sources are then prioritised depending on the potential impact they might

have [100].

Most of the qualitative tools have been developed in the area of project management. But their benefits can be extended to other fields. In the case of petroleum engineering, Zoveidavianpoor and Jalivani have identified which are the main sources of risk and their effects when an oil field is under production by means of fluid injection [101]. The main sources of possible uncertainty they identified are: air, water and ground contamination, high noise levels and excessive water requirements having a direct effect on the availability of surface water for local communities.

For the case of the present study, the main concern is what factors might affect the estimation of a reliable oil production model. A relevant analysis should indicate what factors might lead to issues on the model performance and estimation. A popular qualitative for tool for this task is the Fishbone diagram. The Fishbone diagram also known as the cause and effect diagram graphically shows what affects a certain variable [102]. The diagram looks like a fish skeleton, where each of the bones represents a component that might affect a certain variable, this variable is represented by the fish's head. The following diagram shows the elements that can affect the model's performance and estimation.

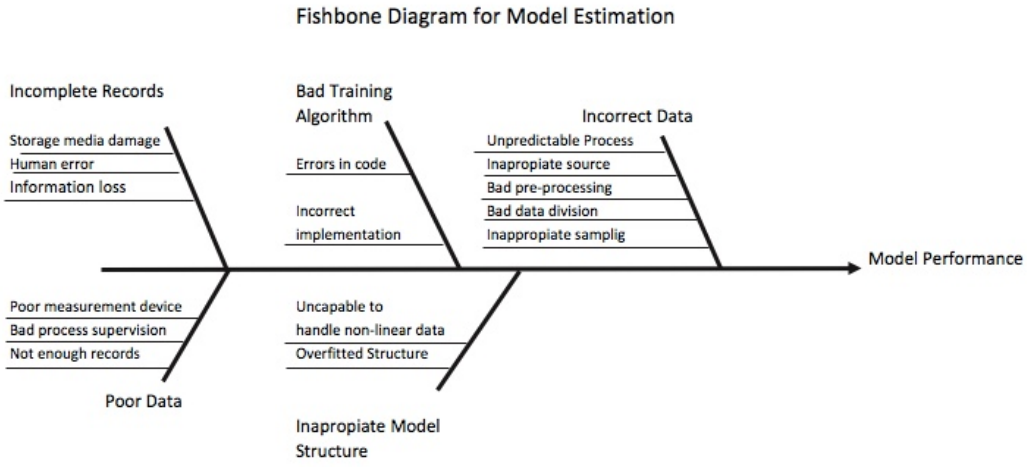

Figure 5.4: Fishbone Diagram

If the estimated model is performing poorly, it is worth looking at the Fishbone diagram to further investigate what might be the root cause of the problem.

5.2.2 Quantitative Methods

Quantitative methods allow us to assign values of occurrence to different scenarios. This makes quantitative methods attractive and reliable as long as the appropriate conditions are considered. According to Comunidad de Madrid [99] the most popular quantitative methods are: Analysis of likelihood, analysis of consequences and computer simulations.

With the development of computational tools, the most popular method in risk analysis are computer simulations using a Monte Carlo strategy. The idea is to use a large number of solutions to evaluate the expectation value of $\varepsilon(f(S_T))$[95]. According to the law of large numbers, if Y_n is a sequence

of random distributed variables, the sequence can be approximated to the following expression:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} Y_j = \varepsilon(Y_1) \tag{5.5}$$

The previous expression shows that using a large number of samples will make an equation converge at a final value. The degree of accuracy will depend on the number of samples. The difficulty comes with selecting the number of samples, the idea is to simulate the future with a sufficiently broad of scenarios to be representative of a real situation.

5.3 Monte Carlo for NARX Models

For the present case study, we used both NARX models, polynomial and single layer neural networks as our risk models, where we applied a Monte Carlo strategy and observed what is the most likely thing to happen in the future.

According to the central limit theorem, when using random samples a function's value can be approximated to the following expression[95]:

$$\varepsilon(X) + \sqrt{\frac{V}{n}} N(0,1) \tag{5.6}$$

Where $\sqrt{\frac{V}{n}}$ is the standard error due to different runs using n number of samples. So in order to find the appropriate value of n, the standard error should not be more than a pre-defined threshold value. It is assumed that the interval from the generated values lies between 0 and 1. The problem with this approach is that we don't know in advance what the expected value is for any of the analysed variables. Some approaches for determining

the optimal number of samples have been developed, but most only give a brief guideline and suggest that sampling depends mainly on the type of phenomenon under study[103] .

Since most oil companies plan their future operations one year in advance, only 12 months ahead of oil production were estimated by using the iterative model's predictive output forecasts.

We designed a simple but effective heuristic methodology for finding an adequate number of iterations. The number of iterations was incremented using different intervals. For every suggested value, one-year forecasts were estimated using the best models from chapters 3 and 4. Only the values of the 12th forecast were analysed since it is on these forecasts were most of the uncertainty happens. If any of the 3 analysed variables (minimum, maximum and mean values) changed more than 10% on a different run, the sampling size would then be increased again, if the variables' value did not change after 5 runs, then the sampling size for the Monte Carlo simulation was determined. A tolerance level of 10% change was used since using a tighter limit would drastically increase the number of required iterations and the obtained benefit would be little. The results from our experiment are summarised on the following tables.

Polynomial Model							
No of Runs	Number of Iterations	Min	Mean	Max	Changes in %		
1	1000	-1.3530	-0.0194	1.1945	Min	Mean	Max
2	1000	-1.7883	-0.0291	1.2057	32.1706	49.7460	0.9372
1	10000	-1.9673	-0.0184	1.4197	NA	NA	NA
2	10000	-1.7830	-0.0190	1.3121	9.3690	3.1879	7.5775
3	10000	-1.8255	-0.0189	1.6033	2.3865	0.0980	22.1928
1	25000	-2.2139	-0.0187	1.7148	NA	NA	NA
2	25000	-2.3963	-0.0237	1.5327	8.2375	26.6713	10.6227
1	50000	-2.3384	-0.0190	1.7751	NA	NA	NA
2	50000	-2.5274	-0.0177	2.0123	8.0830	6.9377	13.3621
3	50000	-2.0047	-0.0205	1.7440	20.6805	15.7315	13.3359
1	75000	-2.0138	-0.0203	1.7440	NA	NA	NA
2	75000	-2.2323	-0.0237	1.5692	10.8531	16.8312	10.0221
1	100000	-2.5561	-0.0214	1.7903	NA	NA	NA
2	100000	-2.4450	-0.0215	1.7880	4.3480	0.7825	0.1285
3	100000	-2.3108	-0.0208	1.8931	5.4882	3.4069	5.8778
4	100000	-2.3022	-0.0228	1.7662	0.3727	9.4679	6.6993
5	100000	-2.3497	-0.0212	1.7559	2.0656	6.6422	0.5861

Table 5.5: Monte Carlo Iterations for Polynomial Model

MSRBF Model							
No of Runs	Number of Iterations	Min	Mean	Max	Changes in %		
1	1000	0.0122	0.0237	0.1972	Min	Mean	Max
2	1000	-0.1419	0.0894	0.3559	1258.1136	277.5240	80.4730
1	10000	-0.1949	0.0296	0.3282	NA	NA	NA
2	10000	-0.2067	0.0903	0.4357	6.0570	204.6200	32.7521
1	25000	-0.2251	0.0499	0.4098	NA	NA	NA
2	25000	-0.2595	0.0902	0.4852	15.3015	80.7120	18.4063
1	50000	-0.2607	0.0566	0.4718	NA	NA	NA
2	50000	-0.2928	0.0903	0.4885	12.3090	59.5343	3.5431
1	75000	-0.3195	0.0680	0.5011	NA	NA	NA
2	75000	-0.3422	0.0902	0.5028	7.0886	32.7479	0.3534
1	100000	-0.3440	0.0735	0.5237	NA	NA	NA
2	100000	-0.3290	0.0904	0.5342	4.3548	22.9181	2.0077
1	150000	-0.3170	0.0679	0.5387	NA	NA	NA
2	150000	-0.3617	0.0903	0.5428	14.1001	33.1041	0.7602
1	200000	-0.3382	0.0736	0.5276	NA	NA	NA
2	200000	-0.3499	0.0904	0.5561	3.4523	22.8028	5.4177
1	250000	-0.3710	0.0969	0.5605	NA	NA	NA
2	250000	-0.3539	0.0904	0.5467	4.6218	6.7112	2.4560
3	250000	-0.3760	0.0903	0.5841	6.2588	0.0392	6.8366
4	250000	-0.3541	0.0903	0.5539	5.8082	0.0469	5.1642
5	250000	-0.3380	0.0904	0.5550	4.5476	0.0479	0.2081

Table 5.6: Monte Carlo Iterations for MSRBF Model

As we can see from the previous tables there is no standard number of iterations for the Monte Carlo simulation. We can also see that polynomial and radial basis function models converge at different rates.

According to our results, the polynomial model requires less than half of the number of iterations (100,000) compared to the radial basis function model (250,000) to converge to a final value for 12 prediction steps ahead.

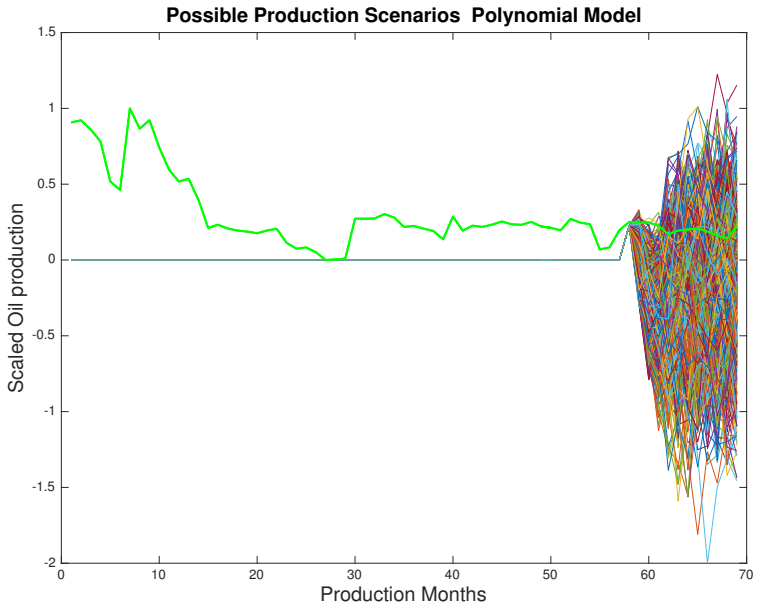

Figure 5.5: Polynomial model predicted scenarios

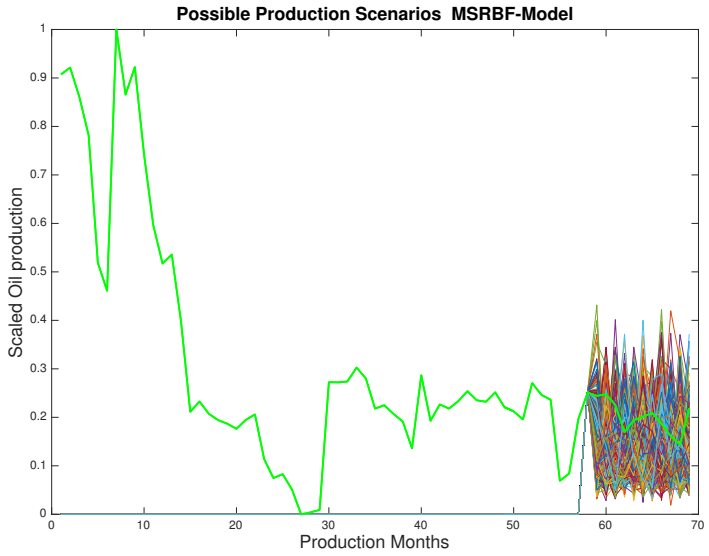

Figure 5.6: MSRBF model predicted scenarios

The previous plots show all the possible future scenarios from the Monte Carlo simulation. Negative output values should be discarded, since it would mean that oil production goes back in to the ground in some cases.

5.3.1 Risk Profile Analysis

As it can be seen on Figures 5.5 and 5.6, the predicted future scenarios from the polynomial and MSRBF NARX models are not the same. So how do we know which model is right and which is not? There is not a yes or no answer since we can only guess what the most likely value will be based on our models.

We have to keep in mind that mathematical models are only an approximation of the studied phenomena and can only partially capture the

dynamics. According to the U.S Energy Information Administration, oil reservoir models should only be considered as tool for finding alternative futures, keeping in mind that future is uncertain and only a small sub-set of the possible scenarios can be modelled under certain conditions [104].

The following plots show how the standard deviation changes with every kth-step ahead the in the predictions, this is a measure of future uncertainty as it exhibits how disperse the data is.

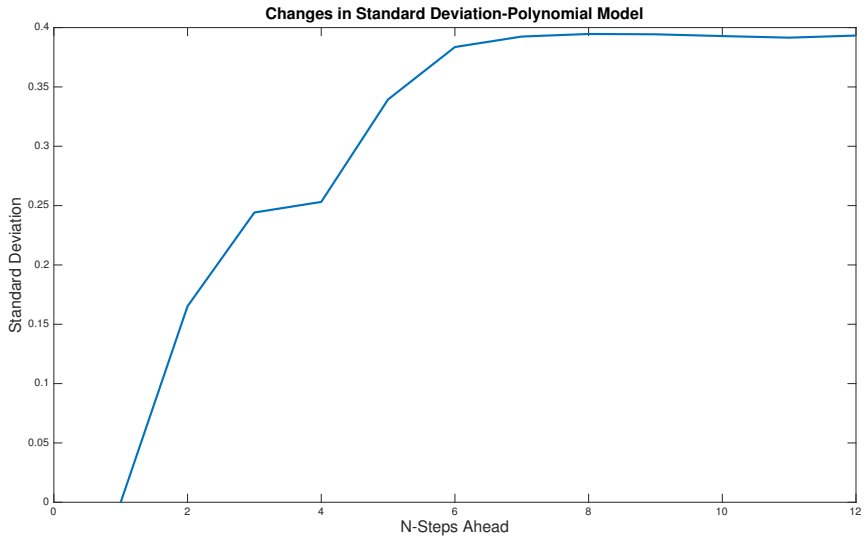

Figure 5.7: Changes in Std Polynomial Model

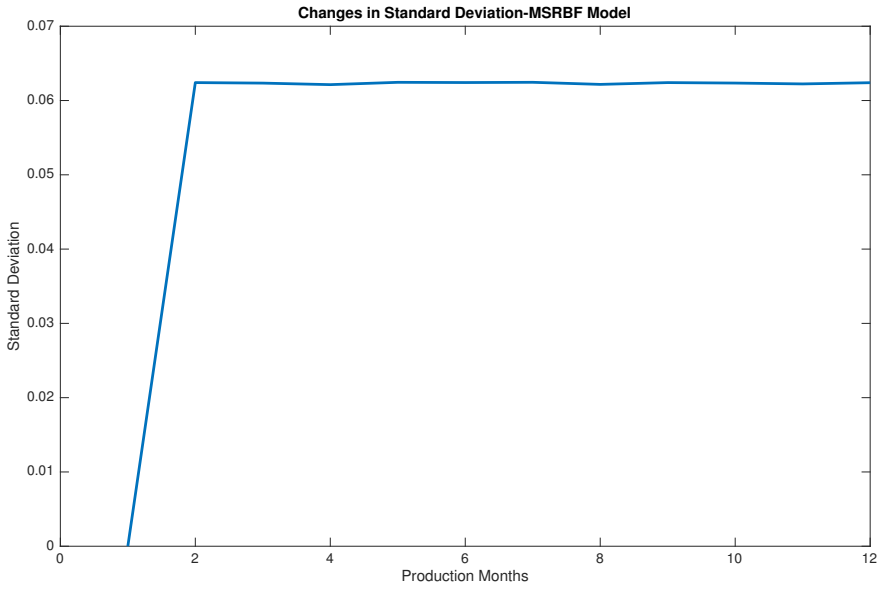

Figure 5.8: Changes in Std MSRBF Model

For the radial basis function model, the standard deviation value remains about the same level for all the future prediction steps. This behaviour agrees with the estimated predictions shown on Figure 5.6 where the values look equally distributed. This is because the maximum values the model can generate are those that are close to the centre's from the radial basis functions. Therefore, any other values will be minimised and filtered out. On the other hand, previous input values have a greater effect on the polynomial model compared to MSRBF forecasts. This behaviour also agrees with Figure 5.5 where data looks more spread as the prediction horizon increases.

Risk analysis theory has developed a useful tool to quantify the uncertainty of a future event. This concept is known as risk profile. A risk profile

is the cumulative probability from $-\infty$ to the value we want to evaluate. If a variable's value is highly uncertain, its Risk Profile will have a wide shape. The variable's uncertainty can be graphically analysed by only looking at the risk profile plot.

There are different distribution functions that can be used for estimating the risk profile, the most popular functions are: Triangular, uniform, discrete and normal distribution [99].

Triangular distribution is very popular, it is very simple and provides a good approximation for many real applications. Uniform distributions are useful when the probability of any of the events is the same. In reality there are few cases where this distribution can be used. Normal distributions are the most used type of approximation for describing the likelihood of an event. Most events on real applications follow a normal distribution and it's been therefore well studied, it is described by the following equation.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-u}{\sigma})^2} \tag{5.7}$$

To determine what distribution was more appropriate for the generated data, a histogram was created giving us a graphical view from the distribution.

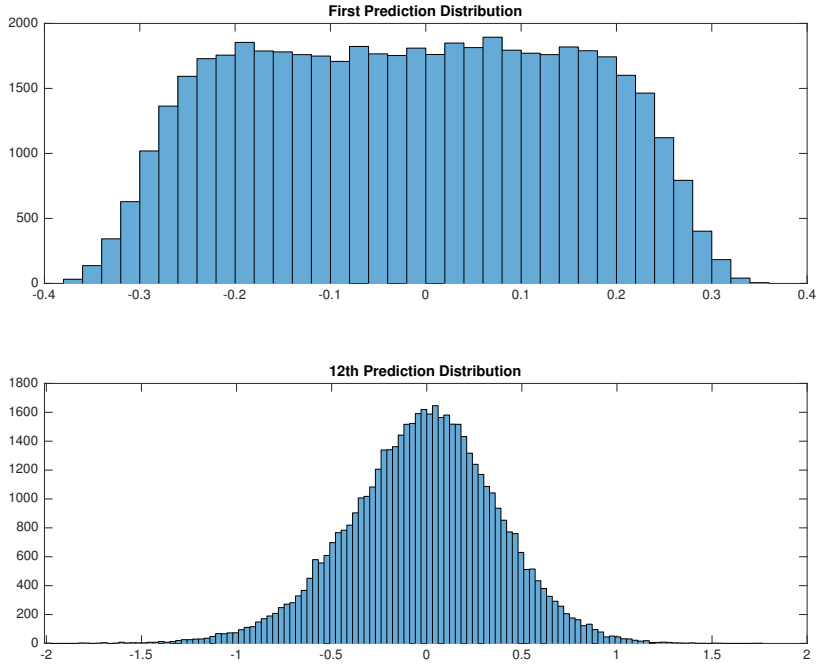

Figure 5.9: Histogram From Polynomial Model

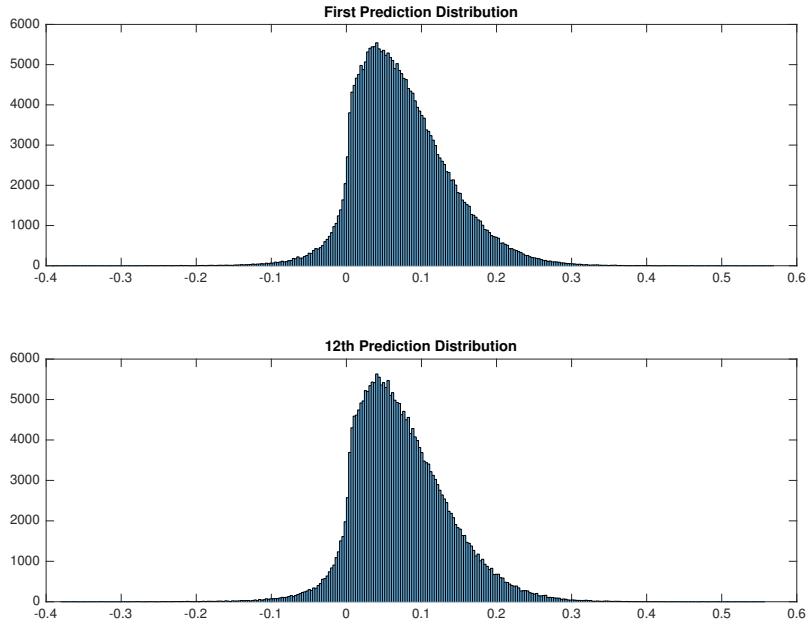

Figure 5.10: Histogram From MSRBF Model

For the polynomial model, it is evident that the predictions follow a normal distribution. As the prediction horizon increases, the forecasts' distribution takes a wider shape. For the first prediction plot, the distribution is very narrow (the histogram has been zoomed in) indicating a high certainty on the corresponding production value.

For the radial basis function model, the distribution from the generated forecasts does not seem to change according to the prediction horizon. According to Figure 5.10 we can see that the distribution is normal, but skewered to the left, this is an indication that more data points lie to the left side of the mean.

After selecting a normal distribution for approximating the distribution of the generated predictions, we could then proceed with estimating the risk profiles for every k-th prediction step ahead. The results can be observed on the following figures.

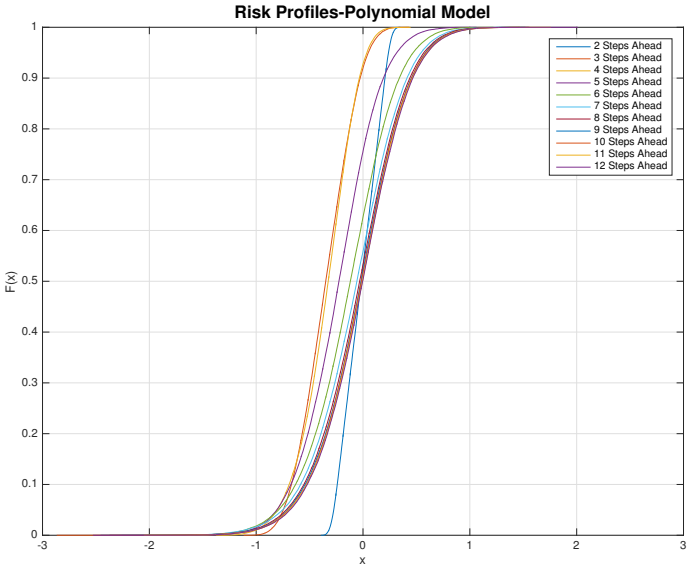

Figure 5.11: Risk Profiles From Polynomial Model

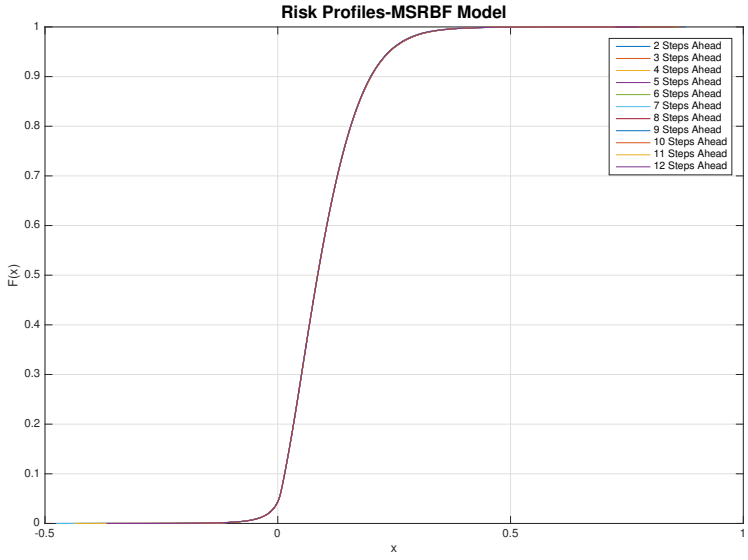

Figure 5.12: Risk Profiles From MSRBF Model

The previous risk profile plots are consistent with the previous analysis. The plots' shape corresponds to a normal distribution, our assumption for using such approximation is then appropriate. For the polynomial model case, the risk profile looks wider as the prediction horizon increases, while for the radial basis function model, the risk profiles overlap each other since the uncertainty remains constant independent of the prediction horizon.

Intuition tells us that uncertainty should indeed increase with the prediction horizon, as every future scenario depends on the previous values leading to an infinite number of possibilities. In this aspect the polynomial model provides a more realistic view into what may happen in the future.

Based on the model's output from the Monte Carlo simulation it is pos-

sible to observe on how many cases the output lies within a certain range, these ranges are known as bins. If the output lies more within certain bins we can say it is more likely for the output to be driven into the range of values specified by those bins. From the Risk profile plot, the horizontal axis represents the normalised out value of the model's output, in our case the normalised oil production value. The vertical axis shows the cumulative probability of the distribution obtained through the Monte Carlo Simulation, in other words the probability of reaching certain production values.

Risk profile plots are very useful, besides graphically showing how uncertain oil production will be in the future months. They provide a simple tool for quantifying how likely it will be to reach the production goals for the company in charge of the operating oil field.

The plot can be used in two different manners. For the first one, if we want to know how likely it is to reach a certain production value at a specific future month, we just have to look for the intersection between our desired production value and the cumulative probability given by the plot without the need of computing any calculations. For the second case, information can be obtained in the opposite direction, if we want to know what production value we can get at a certain probability confidence value, we just have to look for the intersection between our desired probability value and obtain the corresponding production level.

5.4 Conclusions

Determining which are the most important injection wells in the field is a complex task since the relation between injection rates and oil production is non-linear. The polynomial model shows an intuitive scan of which could be the most significant inputs from its equation. Determining the same information through a multi-layer neural network requires the use of a feature selection algorithm which in most cases will be very computationally expensive. It is hard to say which method is more accurate since for industrial applications the true answer can't be known in advance. The methods should be used as complementary tools before making changes on the production process.

When forecasting a future scenario, we always have to consider how likely it is for a certain case to happen. Using Monte Carlo simulations gives us a glimpse of what the future might look like, however we have to consider a few aspects like how far into the future we want to predict, what is the operating input range and the number of iterations we need in order to get a decent guess.

Uncertainty has to do with the distribution of the model's forecast data, the more spread it is, the more uncertain a future scenario will be. Risk profiles developed in mathematical finance are a very useful tool that lets us estimate and visualise how likely it is for a certain scenario to happen at a certain statistical confidence level. According to our analysis and simulations the polynomial model provides a more realistic view into the future, considering that uncertainty on the estimated predictions increases with the pre-

diction horizon, on the other hand the radial basis function model produces estimates within the same range independently to the prediction horizon.

Chapter 6

Production Optimisation

6.1 Problem Definition

Optimisation is defied as the action of making best of most effective use of resources [105]. From this definition we can see that in almost every activity optimally using resources is the ideal goal to reach and must always be implemented.

In the context of the oil and gas industry, the main challenges oil companies face is extracting as much hydrocarbons as possible, at the lowest cost, for the longest period of time. In order to achieve this complex target multiple resources and strategies are required (financial, human and technological are some of the examples in the list) [106].

In the short term, petroleum engineers design a production strategy which will meet every day targets and demand, while reservoir engineers make sure resources are extracted in the most efficient way so that the reser-

voir can produce for a long time while increasing recovery rates [107].

The success of meeting goals and achieving an appropriate production performance depends on the decisions taken to drive future operating conditions to a desirable point. Traditionally, the task has been completed by performing volumetric analysis, material balance and decline curve analysis methods [106]. These methods have proven to be effective but only at limited scale, as reservoirs get more complex and the market is more competitive these methods are no longer sufficient to satisfy the requirements the industry demands. The key to success is then directly related to using the most innovative methods which will lead to taking the best decisions that will then drive operations to the best possible scenario.

Motivated by the current requirements, on this work the latest methods in optimisation were reviewed and implemented.

6.2 Heuristic Optimisation Methods

The problem of optimisation is very important since it provides solutions to issues we deal with in our everyday lives. Many optimisation methods have therefore been developed. The list of available methods includes linear programming, convex problems, multi-objective optimisation, multi disciplinary optimisation, etc. A good review of the most used methodologies in engineering can be found at[108].

The standard optimisation problem can be summarised as follows:

$$\begin{aligned} &Minimise : f(x) \\ g_j(x) &\leq 0 \qquad j = 1, m \\ h_k(x) &= 0 \qquad k = 1, p \\ x_{\mathrm{iL}} \leq x_i &\leq x_{\mathrm{iU}} \qquad i = 1, n \end{aligned} \tag{6.1}$$

From the previous equation, $f(x)$ is the objective function, $g_j(x)$ is an inequality constraint, $h_k(x)$ is an equality constrain function. The independent variable x represents the search space where different values can be used to reach optimality. The search space is limited by x_{iL} and x_{iU}.

Optimisation methods are used to solve the problem stated on Equation 7.1 where a combination of values satisfying the equalities and constrains leading to the best feasible solution is considered as optimal.

The main issues optimisation methodologies have to deal with are: avoiding local optima solutions, time to find solution, ability to handle constrains, applicability in complex problems and solution accuracy.

Originally, optimisation procedures were analytical steps on which, mathematical expressions had to be manipulated to find a solution. As problems increased in complexity and more restrictions on the independent variables had to be considered, the application of such methods became non-trivial and extremely complex. This led to the development of new methodologies.

During the last decade thanks to the advances in computational tools, heuristic optimisation algorithms have gained researchers attention from diverse areas leading to many successful applications. The most popular algorithms can be found on "Handbook of Metaheuristics" [109].

Heuristics are solution methods that look for local improvement solutions and higher level procedures that lead to finding an optimal solution. They are a search procedure that continuously looks for a new solution which is closer to reaching the objective function. These methods have become so effective that they are probably the most used strategies for solving complex problems. Heuristic methods depend on many parameters and the number of solutions evolves as a trial and error process, thus optimality can not be certified, but a very good approximation can be reached which in most cases is enough to satisfy the needs of real application problems [109].

6.3 GRASP Optimisation Using NARX Models

Selecting the best algorithm has become a tricky process since there are many options available, multiple algorithms claim to offer the best solution for a particular problem. This is an issue that is well discussed by Manuel López-Ibañez [110].

The performance of an algorithm can drastically change depending on the tuned parameters, it is therefore really hard to compare the performance of the different options, specially if long time was spent training a particular algorithm. There are different criteria on which the algorithms can be eval-

uated, but depending on the specific problem priorities might be different.

One of the most attractive algorithms is Greedy Randomised Adaptive Search Procedure (GRASP). This algorithm was developed during the mid-nineties by Feo and Resende [111], it has been mainly used in the area of operations research solving issues in: logistics, manufacturing, partition, transport, power systems, telecommunications and biology [112]. According to our literature review, this algorithm has never been used in the area of petroleum engineering solving the problem of maximising future oil recovery by optimising water injection. Due to the success this methodology has had in different areas we decided to combine our developed NARX models with the GRASP optimisation algorithm and solve the problem of maximising oil production on a field under water injection. A review of field injection optimisation using commercial reservoir simulators can be found at [113].

GRASP is attractive compared to other optimisation algorithms since it is simple to implement, it avoids local optimal solutions and its search procedure is intuitive. In some aspects, it may not be the best solution since it can take longer to find a solution compared to other methods, but its proven to give satisfactory results.

The algorithm is a multi-start, iterative metaheuristic procedure consisting of two phases, a global search known as construction and a local search. During the construction phase a solution is built, if the solution is not feasible a new solution is proposed until a feasible one is found. Once a feasible solution has been found the local search looks for a refined solution on the neighbourhood, the best solution is kept as the optimal one [111].

In a greedy algorithm solutions are proposed from scratch within the predefined limits. The higher the number of proposed solutions, the more likely a good solution can be found. The number of proposed solutions is given by a greedy evaluation function. The greediness represents the incremental increase of incorporating a new solution as feasible. In other words, the greedier the algorithm the more precise the final solution would be.

Randomisation in the algorithm plays a very important role, since by randomly sampling within the search space local optima points are avoided. It allows different trajectories to be explored from the initial solutions. The randomisation process requires from the greedy search, since the more random the proposed solutions are, the better the final solution will be.

A neighbourhood solution is defined as a solution with similar or better properties than the global solution found at the construction search. In some occasions neighbourhoods might include infeasible solutions which would then be discarded.

The local search has the objective of finding a refined solution. Solutions are found following an iterative process by restricting the search space within the neighbourhood. The effectiveness of the local search mainly depends on: starting solution, neighbourhood size, cost function and search strategy [111]. Different versions of a GRASP algorithm have been developed in order to handle these issues. The basic version of the GRASP algorithm is composed of the following steps [111]:

1)$Set f^* \longleftarrow \infty$;

2)**For** $k = 1, ..., MaxIterations$ **do**

3)$S \longleftarrow GreedyRandomisedAlgorithm$

4)**if** S is not $feasible$ **then**

5)$S \longleftarrow RepairSolution(S)$

6)**end**;

7)$S \longleftarrow LocalSearch(S)$

8)**if**$f(S) \leq f^*$**then**

9)$S^* \longleftarrow S$;

10)$f^* \longleftarrow (S)$

11)**end**;

12)**end**;

13)**return**S^*;

14)**end**

(6.2)

For the global search, the search space has to be limited within a range, the range is given by C^{min} and C^{max}, where C represents the candidate values of the independent variables to select from. Once feasible solutions have been found, the neighbourhood search space is given by the following expression:

$$\alpha(c^{max} - c^{min}) \tag{6.3}$$

From the previous expression α is a constant value that scales and limits the search space, while "c" represents the range of values where the search

procedure will look for better solutions.

It is important to note that the search algorithm will try to find an optimal solution for each of the independent variables. So the shown steps on Equation 6.2 have to be applied to each independent variable.

Our NARX models are composed of 20 independent variables, which represent the 20 injection wells in the field. The objective is to optimise the injection rates for every future month in order to maximise oil production. The algorithm has to be run according to the following expression:

$$
\begin{aligned}
&1)\textbf{for} \quad i = 1, ..., W(ProductionWells)\\
&2)\textbf{for} \quad k = 1, ..., N(MonthsAhead)\\
&3)RUN \quad GRASP\\
&4)Collect\ best\ input\ set\ from\ u_1, ..., u_M\\
&5)\textbf{end}\\
&6)\textbf{end}
\end{aligned}
\tag{6.4}
$$

From the previous equation, on step 2 we assume that the objective function is to maximise the model's output (oil production). Variable N represents the number of months ahead on which oil production needs to be maximised, while M is the number of inputs on the model.

All the data that is fed into to the models is scaled data, we therefore assume that the injection rates from every well have the possibility to fluc-

tuate from a null value to their previous maximum. The search space is then $[0, 1]$ for every independent variable. In order to obtain the rates at their usual scale, they have to be re-scaled using the following expression:

$$x = x'(max_x - min_x) + min_x \tag{6.5}$$

From the previous expression x is the re-scaled value for the injection rate, x' is the normalised value given by the the search procedure while max_x and min_x are the maximum and minimum values of the original dataset respectively.

To determine what is an appropriate number of iterations on the global search, the approach presented on the previous chapter given by Tables 5.4 and 5.5 was used. It was therefore determined that the polynomial model requires at least 100,000 while the radial basis function needs 250,000 iterations to find an appropriate solution.

One of the advantages of using the GRASP algorithm is that we can know in advance how long it will take to find a solution. If we know how long it takes for one forecast to be estimated, and we know how many possible solutions will be evaluated we simply have to multiply the number of possible solutions by the required time to estimate a forecast as follows:

$$Search\ Time = (1\ Forecast\ Time) * (Number\ of\ Solutions) \tag{6.6}$$

The number of iterations and the size of the local search space are parameters to select before running the algorithm. A suggestion from Silva and colleagues [114] is to limit the local search space to 20% of a solution found

at the construction stage, this means using a value of 0.2 for α from Equation 6.3. The number of proposed solutions for the local search we used was 1500, this number was found in the same manner as the number or iterations for the global search. The main difference was that for the local search, the change in oil production gain was the observed variable. The oil production gain is given by the following equation:

$$Gain = \frac{100}{n}\Big(\sum_{k=1}^{n} yopt_k - \sum_{k=1}^{n} y_k\Big) \tag{6.7}$$

Oil production gain is given as a percentage quantity, n represents the number of steps ahead on which the production has been maximised. The gain is calculated with respect to the measured data. This was done for comparative purposes, if the methodology is to be applied on a real scenario, there would be no measured values to compare with. The procedure would then be to select an approximate guess of what is the most likely scenario in the future or an appropriate reference level and then estimate the change in gain using Equation 6.7.

The following table summarises the results.

No of Local Iterations-Polynomial Model		
Number of Iterations	Max Gain	Gain Change %
50	504	NA
50	612	21.4286
100	603	NA
100	763	26.5340
200	780	NA
200	650	16.6667
300	529	NA
300	697	31.7580
500	536	NA
500	634	18.2836
700	768	NA
700	612	20.3125
1000	866	NA
1000	1100	27.0208
1500	988	NA
1500	1016	2.8340
1500	998	1.7717
1500	1102	10.4208
1500	1083	1.7241

Table 6.1: Number of iterations for local search-Polynomial Model

No of Local Iterations-MSRBF Model		
Number of Iterations	Max Gain	Gain Change %
50	400	NA
50	398	0.5000
50	390	2.0101
50	416	6.6667
50	383	7.9327

Table 6.2: No of iterations for local search-MSRBF Model

The inputs that drive the output to the shown values can be found on Appendix A.

After running the GRASP optimisation algorithm, oil production was maximised by almost 10 times for the polynomial case and 4 times for the radial basis function model.

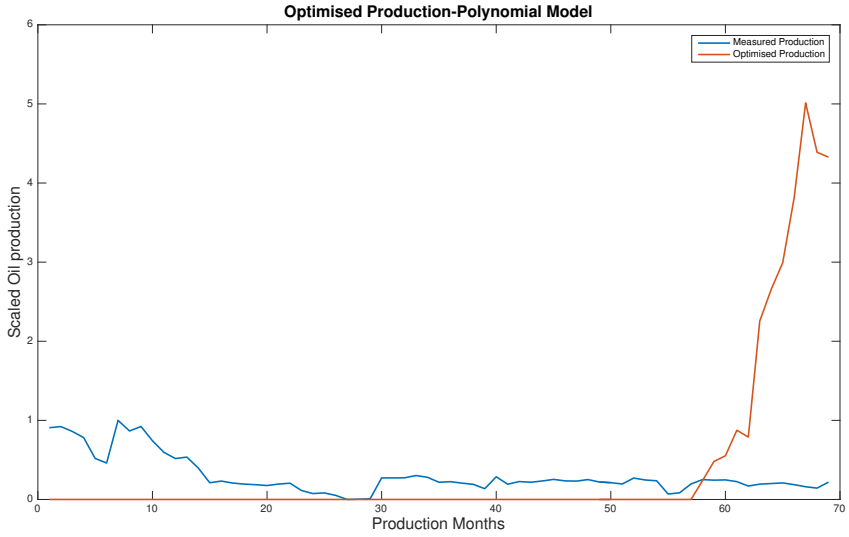

Figure 6.1: Optimised Oil Production-Polynomial Model

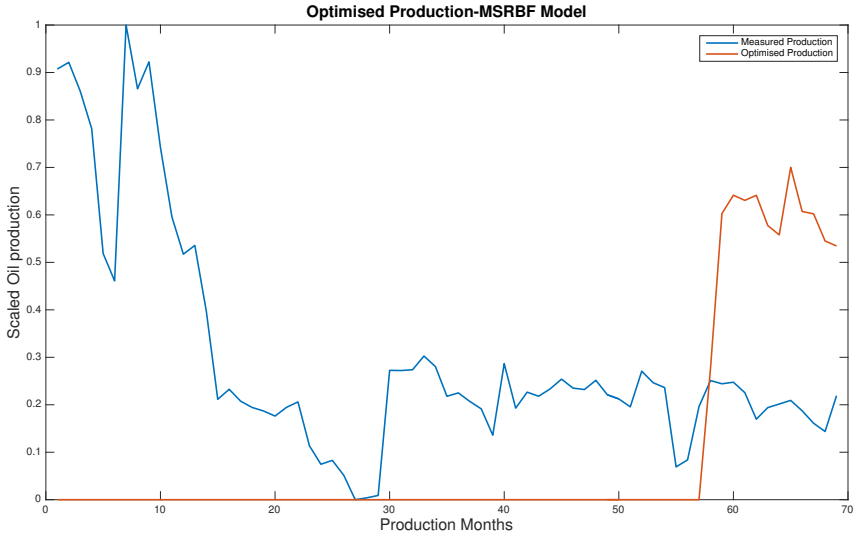

Figure 6.2: Optimised Oil Production-MSRBF Model

As we can see from the previous plots, the extracted oil during the last year of production on the data set is far from the maximised values. The maximised oil production curve is equivalent to selecting the upper limit of the Monte Carlo simulations presented on the previous chapter.

The obtained results seem too good to be true, it seems really unlikely that the production on a well can change so drastically by only changing the injection rates into the field. Making use of the methodology developed on the previous chapter, we can estimate how likely it is to reach a certain production level.

Probability of Reaching Desired Production-Polynomial Model			
Prediction Number	Max Production Value	Probability of Reaching Maximum Production %	Probability of Increasing Production by 30%
1	0.2374	0.00000	0.000
2	0.4805	1.35702	4.892
3	0.5521	0.06063	0.672
4	0.8751	0.10787	0.930
5	0.7892	0.00566	7.450
6	2.2564	0.00570	15.501
7	2.6568	0.00489	19.918
8	2.9922	0.00025	22.221
9	3.8148	0.00033	23.326
10	5.0123	0.00027	23.696
11	4.3888	0.00020	23.912
12	4.3271	0.00084	24.007

Table 6.3: Probability of Reaching Production Levels-Polynomial Model

Probability of Reaching Desired Production-MSRBF Model			
Prediction Number	Max Production Value	Probability of Reaching Maximum Production %	Probability of Increasing Production by 30%
1	0.275800044	0.0000E+00	0.0000
2	0.602853605	0.0000E+00	0.1808
3	0.641459471	0.0000E+00	0.1750
4	0.630595306	0.0000E+00	0.1765
5	0.641290012	0.0000E+00	0.1784
6	0.5775944	3.3307E-14	0.1782
7	0.557867907	4.1078E-13	0.1789
8	0.700181637	0.0000E+00	0.1830
9	0.607136344	0.0000E+00	0.1790
10	0.602248739	0.0000E+00	0.1814
11	0.545134993	1.6875E-12	0.1736
12	0.534500896	8.4932E-12	0.1853

Table 6.4: Probability of Reaching Production Levels-MSRBF Model

The probabilities given on the previous tables are given as percentages. As we can see it is extremely unlikely to reach the maximum production values. It is therefore not recommended to try to obtain the maximum values given by the optimisation algorithm, it involves high uncertainty, it requires higher effort on the injectors and we might end up with little gain. A more appropriate approach would be to try to reach a more conservative rate of production. For example, in literature we have found cases were 30% increase in production was achieved by only modifying the injection rates in the field [115].

The probability of reaching this increase is also given on Tables 6.3 & 6.4. As we can see from the results, according to the forecasts from the polynomial model, increasing the well's production by 30% is possible but really hard to achieve within one year, we can see that the probability of reaching the goal increases with the prediction horizon which intuitively sounds right. In order to be more confident that the desired production goals will be met, there are two options: one is to reduce the objective goals or increasing the time on which the levels will be reached.

According to the radial basis function model the desired production levels are almost impossible to reach within one year this is because the model predicts that the most likely thing to happen is a decrease in production, therefore all increments in production will have a very low probability value. In fact, from the measured data it can be seen that after the 58th measurement production does decrease.

The GRASP optimisation algorithm can be then used to reduce the error between a set point value and the reservoir's oil production output. By using this approach, the optimisation algorithm can be used as a non-linear controller to achieve a certain production value. For example, if we want to increase the production by 30%. We can estimate what the injection values should be to reach this goal. In our case the last scaled production value is 0.1963 so our set point if 30% of the production needed be increased would be 0.2551. To find the required injection rates, the same number of iterations and α value from the maximisation case were used.

The following plots show how the GRASP algorithm was able to keep oil

production at a steady value for both models.

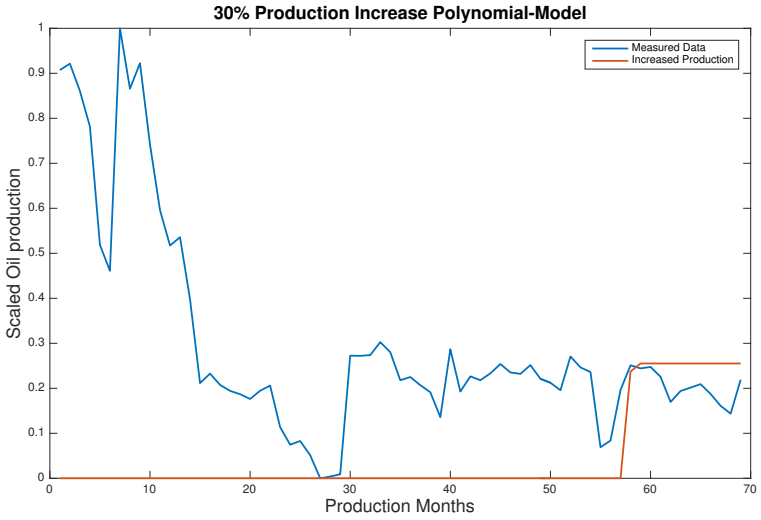

Figure 6.3: 30% Production Increase Polynomial Model

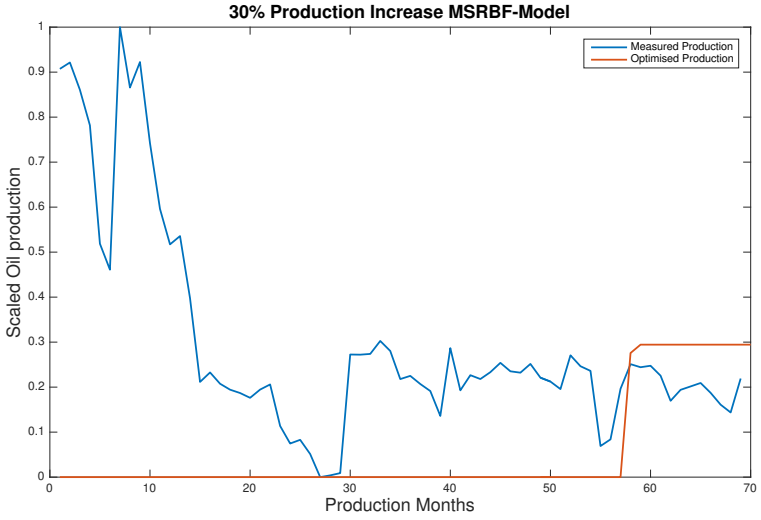

Figure 6.4: 30% Production Increase MSRBF Model

6.3.1 Financial Benefits

The principal attractiveness of our proposed methodology is that accessing the historical records does not represent any extra costs. Operating companies always keep records of the injection/production values in order to assess the field's performance. The analysis does not require anything but a decent personal computer which cost is relatively low (less than £2000). The required investment is therefore minimal compared to the benefits.

Going back to the example of Scott in the North Sea, at its 57th production month, oil production had a value of 334,615m^3 (2104665 *bbl*) see Figure 3.5. On average, oil production in the field had a volume of 431,341m^3 (2,713,053.27 *bbl*).

According to the Nastaq index, Brent crude oil was traded at around $42USD for most of 2016 [116]. The value of the extracted crude oil can be estimated as:

$$P = V * C \tag{6.8}$$

Where P is the total value, V is the sales volume while C is the price at which the oil is being traded. Based on this equation and the current the value of crude Brent oil, the value of the extracted oil production from Scott during the last year was on average $113.9481 Million USD a month.

The following table shows the financial increase compared to the measured production values.

Base Reference	Increase in %	Monthly Increase M-USD	Year Increase M-USD
$113.9481M-USD	2	2.2790	27.3475
Barrel Price	4	4.5579	54.6951
$42 USD	6	6.8369	82.0426
	8	9.1158	109.3901
	10	11.3948	136.7377
	12	13.6738	164.0852
	14	15.9527	191.4327
	16	18.2317	218.7803
	18	20.5107	246.1278
	20	22.7896	273.4753
	22	25.0686	300.8229
	24	27.3475	328.1704
	26	29.6265	355.5179
	28	31.9055	382.8655
	30	34.1844	410.2130

Table 6.5: Financial Increase in Sales

As it can be appreciated on the the previous table, the estimated financial rewards are huge considering that no further equipment or investment has to be made to increase oil production. Even a conservative goal of increasing production by 8% with current low oil prices could represent a gain in more than $100 Million USD a year. Analysing data and taking smart decisions is definitely a strategy that must be considered when increasing recovery rates from a mature field under fluid injection.

6.4 Forecasting Using Ensemble Modelling

Changing the injection rates on an operating oil field can have a great effect on the production performance. We therefore need to be sure that the best injection strategy leading to a promising scenario is applied. If a wrong strategy is chosen, the company's profits can easily be affected as well as leading to low recovery rates.

The decisions based on the generated models are of great importance. But as it has been shown, models are only an approximation of reality and sometimes seem to disagree between each other. The challenge is then to know which model is predicting a truth scenario, based on that the best decisions can be taken [104].

There is no way to know what the future will look like, we can only guess what is the most likely scenario will be. On real applications forecasting is never based on one single model. It's been proved that the combination of predictions coming from different models is on average better than the predictions from any single model.

A good comparison would be the decision taken by an expert on a certain topic compared to the decision taken a a group of experts, the mistakes made by some will be spotted and corrected by the other members on the team.

When the used models are trained using different structures the ensemble is hybrid. When different models were trained using the same algorithm

but with different data the ensemble is non-hybrid [29]. In our case the ensemble is hybrid since the polynomial and radial basis NARX models have completely different structures. In our case, the ensemble predictions are estimated as follows:

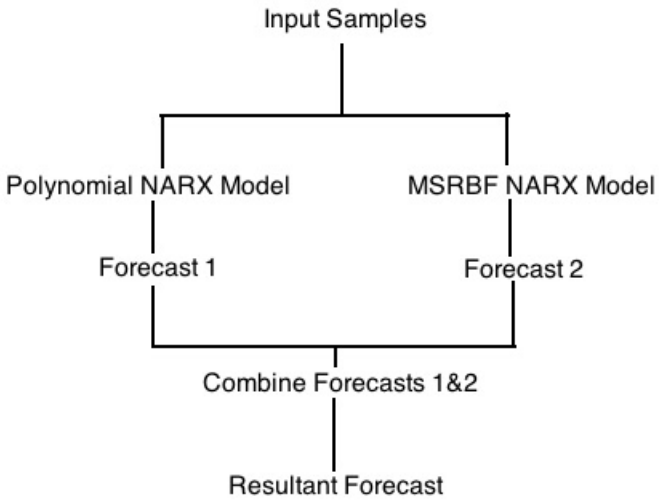

Figure 6.5: Ensemble Diagram

From the previous diagram it can be observed that each model estimates their own forecasts and then the average of both is computed.

The following plots show the performance of the ensemble prediction using the "best" models from chapters 3 and 4.

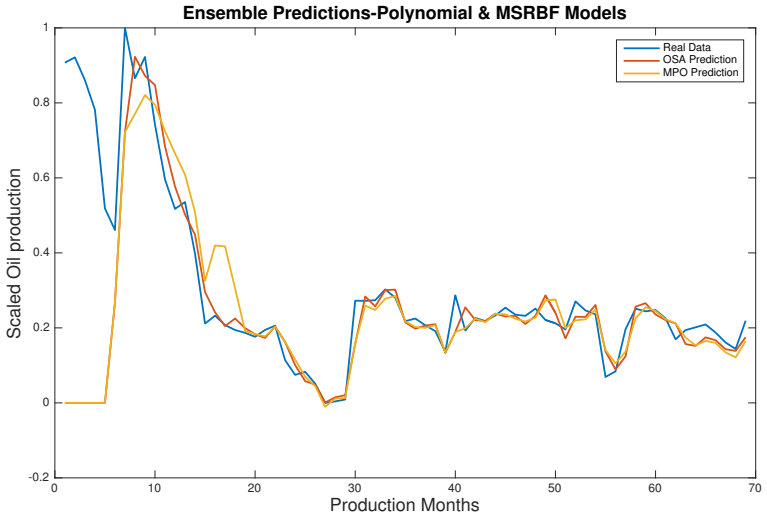

Figure 6.6: Ensemble Forecast Polynomial & MSRBF Model

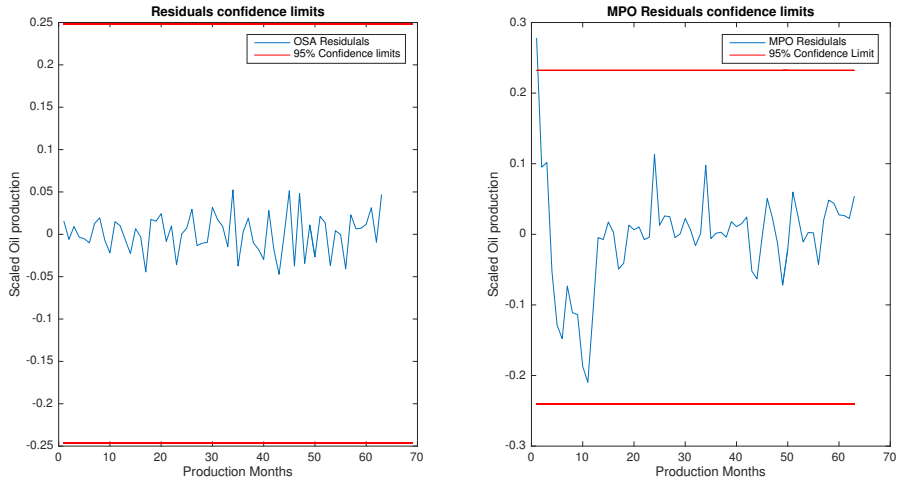

Figure 6.7: Ensemble Residual Limits Polynomial & MSRBF Model

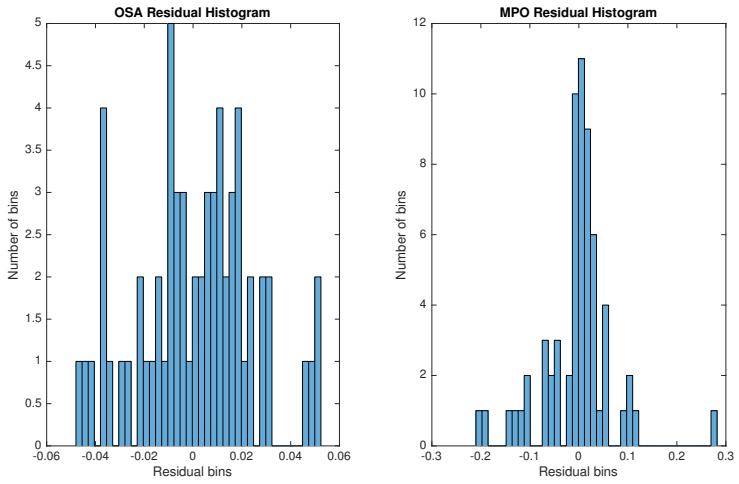

Figure 6.8: Ensemble Prediction Distribution Polynomial & MSRBF Model

As it can be observed on the previous plots, the predictions generated by the ensemble between the polynomial and the radial basis function perform very well. All the residuals lie within the 95% confidence limits and follow a normal distribution. The performance indices are summarised on the following table.

OSA Fit	MPO Fit	MSE OSA	MSE MPO
0.9916	1.0336	0.0001	0.0010
Mean Residuals OSA	Std Residuals OSA	Mean Residuals MPO	Std MPO
0.0010	0.0239	-0.0039	0.0720

Table 6.6: Ensemble Performance Indices

The generated ensemble predictions provide good quality forecasts. This approach will compensate the mistakes individual models might produce.
In terms of keeping the output value at a fixed level, the same approach

should be used where a set of inputs is found for keeping the estimated ensemble output at a steady level. To compute these set of inputs, the GRASP optimisation algorithm was also implemented as shown on the following diagram.

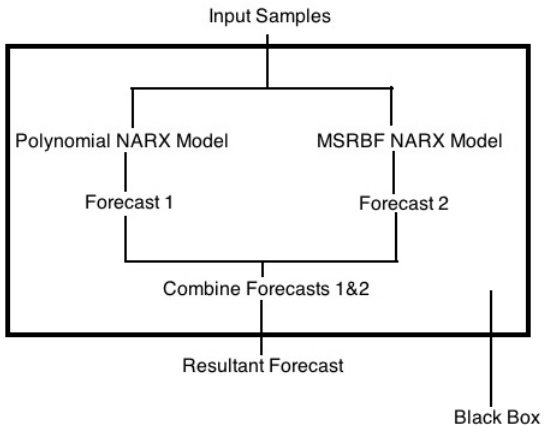

Figure 6.9: Ensemble Modelling Optimisation

The GRASP algorithm can be used in the same manner as it has been shown before but instead of computing estimates through each individual model, a black box model should be used in order to estimate a final forecast. This black box model is the combination of the polynomial and radial basis function NARX models. From the point of view of the optimisation process nothing changes, a set of many random inputs is used and the system's output is observed, the inputs that drive the output to the desired level are selected as the best inputs. On the other hand, the time to compute the estimates will increase since we have to use at least the minimum number of iterations that are required by each individual model. The transparency and interpretation of the model's equations is lost when we apply ensemble

modelling since its very hard to intuitively find the relation between inputs and outputs.

6.5 Conclusions

As it has been shown on the previous examples, the GRASP optimisation algorithm can be used in conjunction with our developed NARX models. The heuristic optimisation algorithm can be used as an estimator to keep the system's output at a certain desired level.

GRASP offers a very intuitive optimisation approach, it avoids local optimal solutions, it is easy to implement and its running time can be estimated in advance. In literature, there are many other optimisation methods, some might claim to be better on different parameters, but for the problem of maximising oil production by means of water injection, any algorithm should not be used to maximise the model's output since it is statistically very unlikely that those values will be reached. The optimisation algorithms should then be used to increase the model's output at a statistically reachable level, considering this, there is no need to use a complex optimisation method that will only estimate an even higher output value since the field's production won't be able to reach such level. In order to reduce uncertainty ensemble forecasting provides an alternative for compensating the biased estimates that individual models might produce.

Chapter 7

Conclusions and Future Work

7.1 Summary

On the current document the development of mathematical tools for the analysis of historical injection/production rates from an operating oil field is presented. The research work is intended to give further understanding for the process of oil production by means of fluid injection. This enables the possibility of increasing recovery rates while optimising resources.

On the first two chapters a review of the traditional and recently developed data based methods for estimating future oil production is presented. Overall, conventional methods provide a general explanation of how oil, gas and water from different sources interact inside the reservoir and how pressure changes affect production. They are good for giving a broad approximation of what the future production values might be and have limited capability when estimating complex what-if scenarios which are necessary when designing an appropriate injection strategy. Due to the requirements

of the modern oil and gas industry these methods are now only used as a complementary tool.

With the increase of computational power from the last decade, researchers have developed data based methods for analysing historical production/injection records. These methods offer an alternative to the traditional reservoir characterisation techniques, but no current method can deal with all the challenges, resulting in either hard implementation for incomplete records, limited long term forecasting capability or black box models which are hard to be linked to the physical parameters on the field.

Since the introduction of the NARMAX methodology by Billings during the late eighties [66], the technique has been successfully applied in different fields of science and engineering but never for Enhanced Oil Recovery modelling. EOR modelling is a complex problem, time delays must be considered while the number of inputs can be significantly large leading to a great number of variables. Chapters 3 and 4 present how to estimate and validate customised polynomial and multi-scale radial basis function NARX models for EOR characterisation. The developed models offer transparent equations which can be easily linked to the physical production parameters, explicitly showing how long it takes for water injection to contribute towards oil production and which are the most significant injection wells. The resulting equations provide valuable information which can be interpreted by non-experts in the field, based on this interpretation high-impact decisions can be made.

Besides providing an interpretable description of the field under opera-

tion, the main purpose of the generated models is to predict future production behaviour, enabling the possibility to run diverse tests without actually implementing them. As a result, there is a reduction in cost and time while avoiding potential undesired scenarios. In order to estimate the likeliness of a future state, a risk analysis methodology based on a Monte Carlo strategy is presented on chapter 5. The presented methodology and concepts are borrowed from financial mathematics and implemented using the generated NARX models. This implementation provides a glimpse of the probability distribution covering future production values.

The main objective of this thesis is to provide a tool for increasing recovery factors on operating fields under fluid injection. To achieve such goal, it is required to implement an optimisation methodology. Chapter 6 provides a survey of the traditional optimisation algorithms comparing them to innovative heuristic optimisation approaches, these methods are a lot easier to implement for complex functions like EOR optimisation. The last chapter shows how to implement a GRASP optimisation algorithm for the purpose of designing an adequate future injection strategy which will maximise production values. The list of available heuristic optimisation algorithms is quite large, but GRASP offers an intuitive approach, avoids local optimal solutions and is easy to implement. The presented study shows that regardless of the selected optimisation algorithm, the results must be combined with a risk analysis in order to design a realistic injection strategy. Failing to do so would lead to unreachable production values and the whole optimisation implementation would be useless.

7.2 Conclusions

The data driven methods developed on this thesis have proven to be capable for modelling active oil fields under fluid injection. The new alternative tool offers understating of the field properties, good prediction performance and a low cost analysis. On the other hand, the new modelling instrument for optimising future injection has not been tested on a real field. This is a necessary step for validating the method's applicability.

The analysis tool developed on this thesis can be summarised on the following steps.

1. Historical Data Pre-Processing.
2. NARMAX Methodology Implementation.
3. Model Structure Analysis.
4. Definition of Future Production Goal.
5. Risk Analysis for Production Goal.
6. Estimation of Future Injection Strategy.

The expected financial benefits of applying the developed methodology are very significant. Every percentage increment in production results in extra profits in the order of millions of dollars. Considering that no further investments have to be made, injecting water at the appropriate rates on the right wells makes the developed data analysis proposal extremely attractive.

7.3 Future Work

The developed methodology for increasing oil production by injecting the optimal amount of fluid into the reservoir is shown on this thesis. However, the presented results and implementation only consider the output of a single production well. In reality, it is required to analyse all production wells within the field. This would require the estimation of either multiple MISO models, one for every production well or a Multiple-Input Multiple-Output (MIMO) model considering all production wells on a single model. Based on these models the same methodology as the one presented on the previous chapters can be extended resulting in the estimation of the best injection strategy for the whole field.

The current studies only show how to increase oil production, but crude oil is not the only fluid that is extracted from the ground on everyday operations. The effect of excessive water or gas extraction towards future production can be very significant. For example, a sudden increase in water production may lead to the effect known as coning where water coming from a close-by aquifer may eventually predominate production. This would require the closure of the producing well leading to a loss in oil production.

A rapid increase in gas production would result in an also rapid pressure drop within the reservoir. This is an undesirable behaviour since a pressure drop would lead to lower oil flow from the ground to the production wells. To compensate for the pressure loss, higher injection rates would be needed, meaning higher operational costs. It is therefore required to predict when these undesirable situations may occur. One approach would be estimating

models for water and gas production in the same manner as the oil production models. Based on these extra models, the future injection strategy would have to consider keeping water and gas production levels at a safe rate besides maximising oil production.

Appendices

Appendix A

MSRBF Model Structure

The following table shows the structure of a multi-scale radial basis function model using one output, 20 inputs, 5 output delays and 4 input delays. The model should be interpreted as on Equation 4.8. For format convenience, input centres and input standard deviation values have been transposed. The table showing input centres should be 6x80 (6 Terms, 20 Inputs by 4 Delays per input), and the standard deviation table should be 6x20 (6 Terms, 20 Inputs).

Theta	Centres Y				
3.1014	0.4610	0.5186	0.7817	0.8607	0.9214
-2.7986	0.4610	0.5186	0.7817	0.8607	0.9214
0.2782	0.2179	0.2803	0.3026	0.2738	0.2721
2.0236	0.5960	0.7421	0.9224	0.8658	1.0000
-75.7711	0.5960	0.7421	0.9224	0.8658	1.0000
0.2186	0.2210	0.2516	0.2321	0.2352	0.2540

Table A.1: MSRBF Model Ny=5 Nu=4

Centres U					
0.73069	0.73069	0.62806	0.87696	0.87696	0.64705
0.63988	0.63988	0.70602	1.00000	1.00000	0.61978
0.61622	0.61622	0.78788	0.25883	0.25883	0.75351
0.00000	0.00000	0.62152	0.88519	0.88519	0.46365
0.84632	0.84632	0.42457	0.70000	0.70000	0.48186
0.82282	0.82282	0.46435	0.78303	0.78303	0.44906
0.79258	0.79258	0.49059	0.77153	0.77153	0.53811
0.00000	0.00000	0.40000	0.71432	0.71432	0.58120
0.73368	0.73368	0.25048	0.36854	0.36854	0.16158
0.67880	0.67880	0.28717	0.14736	0.14736	0.43521
0.78052	0.78052	0.26444	0.52471	0.52471	0.00000
0.84298	0.84298	0.26052	0.49811	0.49811	0.15867
0.00000	0.00000	0.04623	0.00000	0.00000	0.36466
0.00000	0.00000	0.25809	0.00000	0.00000	0.39022
0.00000	0.00000	0.39721	0.00000	0.00000	0.59861
0.00165	0.00165	0.35455	0.00000	0.00000	0.46507
0.00000	0.00000	0.76276	0.00000	0.00000	0.45284
0.00000	0.00000	0.68013	0.00000	0.00000	0.42241
0.00000	0.00000	0.56187	0.00000	0.00000	0.49519
0.00000	0.00000	0.50423	0.00000	0.00000	0.35915
0.00000	0.00000	0.82953	0.00000	0.00000	0.56042
0.00000	0.00000	0.83963	0.00000	0.00000	0.59083
0.00000	0.00000	0.76642	0.00000	0.00000	0.74500

Table A.2: MSRBF Model Ny=5 Nu=4

Centres U					
0.00000	0.00000	0.65934	0.00000	0.00000	0.52998
0.48536	0.48536	0.26187	0.58243	0.58243	0.53352
0.54315	0.54315	0.57091	0.64688	0.64688	0.00000
0.50672	0.50672	0.59372	0.58959	0.58959	0.59158
0.62605	0.62605	0.47743	0.63396	0.63396	0.58402
0.47635	0.47635	0.30662	0.67701	0.67701	0.79943
0.60969	0.60969	0.71027	0.74092	0.74092	0.00000
0.57454	0.57454	0.71420	0.59300	0.59300	0.73994
0.67057	0.67057	0.54319	0.65993	0.65993	0.77093
0.00000	0.00000	0.85330	0.31976	0.31976	0.69489
0.00000	0.00000	0.86691	0.00088	0.00088	0.00000
0.00000	0.00000	0.88048	0.09077	0.09077	0.60212
0.00000	0.00000	0.68712	0.00000	0.00000	0.28945
0.00000	0.00000	0.85035	0.00000	0.00000	0.83074
0.00000	0.00000	0.81477	0.00000	0.00000	0.00000
0.00000	0.00000	0.74279	0.00000	0.00000	0.88084
0.00000	0.00000	0.58220	0.00000	0.00000	0.94296
0.00000	0.00000	0.57622	0.00000	0.00000	0.65821
0.00000	0.00000	0.89939	0.00000	0.00000	0.00000
0.00000	0.00000	0.87664	0.00000	0.00000	0.60013
0.00000	0.00000	0.70606	0.00000	0.00000	0.49985
0.89728	0.89728	0.80670	0.51457	0.51457	0.76597
0.65332	0.65332	0.99907	0.55910	0.55910	0.74817
0.59370	0.59370	0.90012	0.67377	0.67377	0.77264

Table A.3: MSRBF Model Ny=5 Nu=4

Centres U					
0.74880	0.74880	0.61306	0.65841	0.65841	0.71230
0.21349	0.21349	0.01872	0.84547	0.84547	0.00000
0.38271	0.38271	0.32364	0.91267	0.91267	0.00000
0.30277	0.30277	0.47300	1.00000	1.00000	0.00000
0.00000	0.00000	0.32538	0.90117	0.90117	0.00000
0.46762	0.46762	0.81929	0.75603	0.75603	0.39281
0.66881	0.66881	0.74987	0.82437	0.82437	0.38383
0.62291	0.62291	0.76979	0.66684	0.66684	0.39978
0.73142	0.73142	0.54279	0.76916	0.76916	0.42397
0.41037	0.41037	0.55777	0.56649	0.56649	0.00000
0.63736	0.63736	0.74858	0.59582	0.59582	0.00000
0.59465	0.59465	0.69438	0.61847	0.61847	0.29826
0.00000	0.00000	0.44052	0.56584	0.56584	1.00000
0.53059	0.53059	0.76854	0.65760	0.65760	0.00682
0.83423	0.83423	0.83874	0.71420	0.71420	0.29144
0.73440	0.73440	0.89836	0.78959	0.78959	0.71685
0.85329	0.85329	0.64104	0.57718	0.57718	0.45122
0.00000	0.00000	0.71138	0.00000	0.00000	1.00000
0.00000	0.00000	0.79014	0.00000	0.00000	0.99147
0.00000	0.00000	0.70138	0.00000	0.00000	0.97211
0.00000	0.00000	0.41261	0.00000	0.00000	0.64038

Table A.4: MSRBF Model Ny=5 Nu=4

Centres U					
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.37793	0.00000	0.00000	0.23880
0.00000	0.00000	0.01510	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	0.31936
0.00000	0.00000	0.00000	0.00000	0.00000	0.35712
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Table A.5: MSRBF Model Ny=5 Nu=4

Std Y
1.58114
1.58114
0.39528
0.39528
0.19764
0.19764

Table A.6: MSRBF Model Ny=5 Nu=4

Std U					
3.16228	1.58114	0.79057	3.16228	0.79057	1.58114
3.16228	3.16228	0.79057	3.16228	0.79057	3.16228
3.16228	3.16228	0.79057	3.16228	0.79057	3.16228
3.16228	3.16228	0.79057	3.16228	0.79057	3.16228
3.16228	3.16228	0.79057	3.16228	0.79057	3.16228
3.16228	3.16228	0.79057	3.16228	0.79057	3.16228
3.16228	3.16228	1.58114	3.16228	0.79057	3.16228
3.16228	3.16228	3.16228	3.16228	0.79057	3.16228
3.16228	3.16228	3.16228	3.16228	0.79057	3.16228
3.16228	3.16228	3.16228	3.16228	0.79057	3.16228
3.16228	3.16228	3.16228	3.16228	0.79057	3.16228
3.16228	3.16228	3.16228	3.16228	0.79057	3.16228
3.16228	3.16228	3.16228	3.16228	1.58114	3.16228
3.16228	3.16228	3.16228	3.16228	3.16228	3.16228
3.16228	3.16228	3.16228	3.16228	3.16228	3.16228
3.16228	3.16228	3.16228	3.16228	3.16228	3.16228
3.16228	3.16228	3.16228	3.16228	3.16228	3.16228
3.16228	3.16228	3.16228	3.16228	3.16228	3.16228
3.16228	3.16228	3.16228	3.16228	3.16228	3.16228
3.16228	3.16228	3.16228	3.16228	3.16228	3.16228

Table A.7: MSRBF Model Ny=5 Nu=4

A.1 Multi-Layer Neural Network Performance

Model	Delays	Hidden Neurons	MSE Val	MSE Training
1	1	30	1.32E-03	1.25E-02
2	2	30	5.44E-02	8.54E-06
3	3	30	2.22E-02	6.84E-04
4	4	30	7.20E-02	9.14E-04
5	5	30	2.66E-02	1.33E-04
6	6	30	3.07E-02	6.69E-24
7	7	30	2.02E-02	1.07E-04
8	8	30	1.96E-02	1.03E-02
9	9	30	7.27E-02	1.12E-18
10	10	30	1.89E-02	5.71E-03
11	11	30	4.35E-03	3.69E-20
12	12	30	5.16E-03	1.45E-04

Table A.8: 30-Hidden Neurons Model

Model	Delays	Hidden Neurons	MSE Val	MSE Training
1	1	21	1.13E-02	1.35E-04
2	2	21	6.39E-02	3.31E-21
3	3	21	1.16E-04	1.40E-02
4	4	21	2.16E-02	7.10E-05
5	5	21	6.56E-02	2.16E-06
6	6	21	4.24E-02	3.46E-14
7	7	21	2.24E-01	7.59E-03
8	8	21	2.56E-02	1.72E-04
9	9	21	2.11E-02	4.79E-05
10	10	21	1.59E-01	2.55E-02
11	11	21	3.73E-02	2.80E-01
12	12	21	1.78E-05	1.44E-02

Table A.9: 21-Hidden Neurons Model

Model	Delays	Hidden Neurons	MSE Val	MSE Training
1	1	40	1.94E-02	1.88E-02
2	2	40	1.02E-01	1.67E-05
3	3	40	7.27E-02	5.30E-04
4	4	40	3.92E-02	6.41E-21
5	5	40	2.20E-05	1.26E-01
6	6	40	1.74E-02	7.27E-02
7	7	40	5.13E-02	5.86E-12
8	8	40	8.16E-02	1.21E-09
9	9	40	5.98E-02	4.42E-08
10	10	40	1.20E-01	3.67E-12
11	11	40	3.62E-01	9.55E-02
12	12	40	6.48E-02	1.92E-03

Table A.10: 40-Hidden Neurons Model

Appendix B

Production Well Models

The following tables show the equation models for all production wells and the net ranking of the injection wells as explained on section 5.1. There are no models for wells 11, 14,15,17,21 since their corresponding records are too poor to build a model.

Production well 1	Injection Well	Net Ranking
'y(k-1)'	1	3
'u(k-1,3)*y(k-5)'	2	0
'u(k-3,3)*y(k-4)'	3	1
'u(k-3,4)*u(k-1,13)'	4	5
'u(k-3,15)*u(k-4,15)'	5	0
'u(k-3,1)*u(k-4,16)'	6	0
'u(k-3,12)*y(k-3)'	7	0
'u(k-3,16)*y(k-1)'	8	10
'u(k-2,11)*u(k-3,15)'	9	8
'u(k-3,9)*y(k-3)'	10	9
'u(k-2,1)*y(k-5)'	11	7
'u(k-5,3)*u(k-3,10)'	12	6
'y(k-2)'	13	4
'u(k-3,1)*u(k-5,8)'	14	0
'u(k-4,13)*u(k-5,13)'	15	2
	16	5
	17	0
	18	0
	19	0
	20	0

Table B.1: Production Well 1

Production well 2	Injection Well	Net Ranking
'y(k-1)'	1	2
'u(k-1,1)*y(k-2)'	2	5
'y(k-2)*y(k-1)'	3	4
'u(k-2,1)*u(k-1,16)'	4	0
'u(k-3,13)*y(k-6)'	5	10
'u(k-1,1)*y(k-1)'	6	0
'u(k-3,13)*u(k-5,19)'	7	0
'u(k-5,2)*u(k-2,3)'	8	0
'u(k-4,16)*u(k-4,16)'	9	8
'u(k-3,10)*u(k-3,13)'	10	7
'u(k-5,3)*u(k-2,11)'	11	8
'u(k-5,3)*u(k-1,9)'	12	0
'u(k-1,5)*u(k-1,13)'	13	1
'u(k-4,2)*u(k-5,14)'	14	11
'u(k-1,3)*u(k-2,9)'	15	0
	16	2
	17	0
	18	0
	19	6
	20	0

Table B.2: Production Well 2

Production well 3	Injection Well	Net Ranking
'y(k-1)'	1	7
'u(k-1,3)*y(k-6)'	2	1
'u(k-1,12)'	3	0
'u(k-2,1)*y(k-1)'	4	9
'u(k-2,13)*u(k-3,19)'	5	0
'u(k-3,2)*u(k-3,2)'	6	0
'u(k-1,13)*u(k-3,19)'	7	10
'u(k-3,2)*y(k-1)'	8	0
'u(k-2,15)*u(k-5,17)'	9	0
'u(k-2,15)*y(k-2)'	10	2
'u(k-1,4)*u(k-4,17)'	11	0
'u(k-2,7)*u(k-1,13)'	12	5
'u(k-1,7)*u(k-1,14)'	13	2
'u(k-4,4)*u(k-4,4)'	14	10
'u(k-4,4)*y(k-1)'	15	5
	16	0
	17	7
	18	0
	19	2
	20	0

Table B.3: Production Well 3

Production well 4	Injection Well	Net Ranking
y(k-1)'	1	5
'u(k-3,1)*u(k-1,4)'	2	0
'u(k-5,9)*y(k-4)'	3	0
'u(k-5,15)*y(k-3)'	4	1
'u(k-2,4)*u(k-5,19)'	5	0
'u(k-2,15)*u(k-4,17)'	6	0
'u(k-3,7)*u(k-1,11)'	7	4
'u(k-2,11)*y(k-6)'	8	9
'u(k-2,11)*u(k-2,15)'	9	6
'u(k-2,4)*u(k-5,7)'	10	0
	11	3
	12	0
	13	0
	14	0
	15	2
	16	0
	17	8
	18	0
	19	7
	20	0

Table B.4: Production Well 4

Production well 5	Injection Well	Net Ranking
'y(k-1)'	1	0
'u(k-1,4)*u(k-5,4)'	2	0
'u(k-5,4)*u(k-2,12)'	3	0
'u(k-2,7)*u(k-3,11)'	4	1
'u(k-4,7)*u(k-4,19)'	5	0
'u(k-1,4)*u(k-4,6)'	6	0
'u(k-4,4)*u(k-1,14)'	7	2
	8	0
	9	0
	10	0
	11	4
	12	3
	13	0
	14	6
	15	0
	16	0
	17	0
	18	0
	19	5
	20	0

Table B.5: Production Well 5

Production well 6	Injection Well	Net Ranking
'y(k-1)'	1	10
'u(k-1,3)*u(k-1,3)'	2	11
'y(k-6)*y(k-4)'	3	1
'u(k-5,15)*y(k-2)'	4	0
'u(k-1,3)*u(k-1,12)'	5	3
'u(k-1,15)*u(k-2,16)'	6	0
'u(k-3,9)*u(k-5,17)'	7	0
'u(k-5,3)*u(k-3,5)'	8	0
'u(k-1,3)*u(k-4,14)'	9	5
'u(k-1,5)*u(k-1,15)'	10	0
'u(k-5,13)*y(k-4)'	11	0
'u(k-4,13)*y(k-1)'	12	4
'u(k-2,5)*u(k-5,13)'	13	7
'u(k-3,1)*u(k-4,3)'	14	9
'u(k-2,2)*u(k-1,9)'	15	2
	16	5
	17	7
	18	0
	19	0
	20	0

Table B.6: Production Well 6

Production well 7	Injection Well	Net Ranking
'y(k-1)'	1	11
'u(k-4,3)*u(k-1,11)'	2	2
'u(k-3,10)*u(k-5,11)'	3	4
'u(k-3,2)*u(k-3,16)'	4	11
'u(k-3,12)*u(k-2,13)'	5	7
'u(k-3,11)*u(k-3,19)'	6	0
'u(k-1,9)*u(k-3,16)'	7	0
'u(k-3,5)*u(k-2,13)'	8	13
'u(k-4,2)*u(k-2,13)'	9	9
'u(k-3,1)*u(k-5,3)'	10	7
'u(k-1,4)*u(k-3,5)'	11	1
'u(k-2,2)*u(k-3,8)'	12	5
'u(k-1,3)*u(k-3,9)'	13	2
'u(k-4,2)*u(k-1,15)'	14	0
'u(k-2,2)*u(k-2,4)'	15	14
	16	5
	17	0
	18	0
	19	10
	20	0

Table B.7: Production Well 7

Production well 8	Injection Well	Net Ranking
'y(k-1)'	1	8
'u(k-1,3)*y(k-2)'	2	2
'u(k-4,2)*u(k-1,3)'	3	1
'u(k-1,18)*y(k-6)'	4	3
'u(k-4,11)*y(k-3)'	5	0
'u(k-4,15)*y(k-2)'	6	0
'u(k-4,1)*y(k-1)'	7	9
'u(k-1,2)*u(k-1,13)'	8	0
'u(k-1,3)*u(k-2,13)'	9	0
'u(k-1,2)*y(k-6)'	10	0
'u(k-1,11)*u(k-3,11)'	11	5
'u(k-1,3)*u(k-1,3)'	12	0
'u(k-1,4)*u(k-3,11)'	13	4
'u(k-1,7)*y(k-1)'	14	0
'u(k-2,17)*y(k-4)'	15	7
	16	0
	17	10
	18	6
	19	0
	20	0

Table B.8: Production Well 8

Production well 9	Injection Well	Net Ranking
'y(k-1)'	1	4
'u(k-4,2)*u(k-1,3)'	2	1
'u(k-5,1)*y(k-1)'	3	2
'u(k-1,3)*u(k-3,7)'	4	7
'u(k-5,3)*u(k-2,15)'	5	11
'u(k-1,13)*u(k-4,14)'	6	0
'u(k-1,18)*u(k-3,19)'	7	9
'u(k-5,1)*u(k-1,13)'	8	0
'u(k-5,2)*u(k-1,3)'	9	0
'u(k-2,4)*u(k-5,16)'	10	0
'u(k-1,3)*u(k-3,15)'	11	0
'u(k-2,2)*u(k-3,15)'	12	0
'u(k-2,5)*y(k-6)'	13	0
'u(k-3,5)*y(k-6)'	14	0
'u(k-3,2)*u(k-4,15)'	15	5
	16	10
	17	7
	18	0
	19	6
	20	0

Table B.9: Production Well 9

Production well 10	Injection Well	Net Ranking
'y(k-1)'	1	0
'u(k-1,3)*u(k-1,3)'	2	2
'u(k-5,2)*y(k-1)'	3	1
'u(k-1,3)*u(k-1,7)'	4	5
'y(k-2)*y(k-1)'	5	7
'u(k-4,13)*y(k-4)'	6	0
'u(k-2,4)*y(k-1)'	7	3
'u(k-3,5)*u(k-1,12)'	8	0
'u(k-3,2)*u(k-5,9)'	9	5
	10	0
	11	0
	12	7
	13	4
	14	0
	15	0
	16	0
	17	0
	18	0
	19	0
	20	0

Table B.10: Production Well 10

Production well 12	Injection Well	Net Ranking
'y(k-1)'	1	0
'u(k-3,3)*u(k-1,6)'	2	0
'u(k-5,11)*u(k-5,16)'	3	3
'u(k-5,6)*u(k-2,9)'	4	0
'u(k-3,11)*u(k-4,14)'	5	0
	6	1
	7	0
	8	0
	9	5
	10	0
	11	2
	12	0
	13	0
	14	6
	15	0
	16	4
	17	0
	18	0
	19	0
	20	0

Table B.11: Production Well 12

Production well 13	Injection Well	Net Ranking
'u(k-2,16)*y(k-1)'	1	0
'u(k-2,18)*u(k-2,19)'	2	0
'u(k-3,13)*u(k-4,17)'	3	0
'u(k-4,11)*u(k-4,13)'	4	0
'u(k-1,11)*y(k-5)'	5	0
	6	0
	7	0
	8	0
	9	0
	10	0
	11	7
	12	6
	13	1
	14	0
	15	0
	16	2
	17	5
	18	3
	19	3
	20	0

Table B.12: Production Well 13

Production well 16	Injection Well	Net Ranking
'y(k-1)'	1	3
'u(k-4,3)*u(k-5,3)'	2	0
'u(k-4,12)*y(k-2)'	3	0
'u(k-3,7)*u(k-5,7)'	4	1
'u(k-4,5)*y(k-5)'	5	4
'u(k-5,4)*u(k-5,4)'	6	0
'u(k-5,4)*y(k-5)'	7	2
'u(k-1,16)*y(k-2)'	8	0
'u(k-5,11)*u(k-3,13)'	9	0
'u(k-4,7)*u(k-1,17)'	10	10
'u(k-1,4)*u(k-1,10)'	11	7
'u(k-3,5)*y(k-4)'	12	5
	13	7
	14	0
	15	0
	16	6
	17	9
	18	0
	19	0
	20	0

Table B.13: Production Well 16

Production well 18	Injection Well	Net Ranking
'y(k-1)'	1	5
'u(k-5,2)*u(k-3,16)'	2	3
'u(k-5,4)*u(k-2,6)'	3	3
'u(k-5,1)*u(k-3,6)'	4	8
'u(k-3,6)*u(k-4,12)'	5	0
'u(k-5,2)*u(k-3,3)'	6	1
'u(k-5,3)*u(k-2,13)'	7	6
'u(k-3,1)*u(k-4,7)'	8	0
'u(k-5,7)*u(k-3,12)'	9	0
'u(k-3,12)*y(k-3)'	10	0
'u(k-1,3)*u(k-5,19)'	11	0
	12	2
	13	9
	14	0
	15	0
	16	7
	17	0
	18	0
	19	10
	20	0

Table B.14: Production Well 18

Production well 19	Injection Well	Net Ranking
'y(k-1)'	1	3
'u(k-5,1)*u(k-5,3)'	2	0
'u(k-1,3)*u(k-1,13)'	3	1
'u(k-3,5)*u(k-5,13)'	4	0
	5	4
	6	0
	7	0
	8	0
	9	0
	10	0
	11	0
	12	0
	13	2
	14	0
	15	0
	16	0
	17	0
	18	0
	19	0
	20	0

Table B.15: Production Well 19

Production well 20	Injection Well	Net Ranking
'u(k-2,7)*y(k-1)'	1	0
'u(k-3,5)*u(k-5,13)'	2	8
'u(k-3,6)*u(k-5,13)'	3	5
'u(k-2,3)*u(k-4,5)'	4	7
'u(k-1,10)*u(k-5,13)'	5	3
'u(k-2,10)*u(k-5,13)'	6	2
'u(k-1,6)*u(k-5,13)'	7	6
'u(k-3,2)*u(k-3,4)'	8	0
'u(k-1,6)*u(k-3,13)'	9	9
'u(k-3,16)*y(k-5)'	10	4
'u(k-3,13)*y(k-6)'	11	0
'u(k-1,3)*u(k-4,4)'	12	0
'u(k-5,9)*u(k-4,13)'	13	1
'u(k-4,6)*u(k-3,13)'	14	0
	15	0
	16	0
	17	0
	18	0
	19	0
	20	0

Table B.16: Production Well 20

Production well 22	Injection Well	Net Ranking
'y(k-1)'	1	0
'u(k-5,5)*u(k-3,14)'	2	6
'u(k-5,6)*u(k-2,13)'	3	0
'u(k-5,4)*u(k-2,13)'	4	1
'u(k-1,2)*y(k-1)'	5	3
'u(k-1,4)*u(k-1,4)'	6	5
'u(k-3,13)*u(k-2,19)'	7	8
'u(k-1,4)*u(k-1,7)'	8	0
	9	0
	10	0
	11	0
	12	0
	13	2
	14	3
	15	0
	16	0
	17	0
	18	0
	19	7
	20	0

Table B.17: Production Well 22

Production well 23	Injection Well	Net Ranking
'u(k-4,2)*u(k-3,17)'	1	0
'u(k-4,4)*u(k-5,13)'	2	3
'u(k-2,4)*u(k-3,15)'	3	2
'u(k-5,4)*u(k-5,13)'	4	1
'u(k-1,3)*y(k-5)'	5	9
'u(k-2,4)*u(k-4,5)'	6	11
'u(k-2,4)*u(k-1,12)'	7	0
'u(k-2,11)*u(k-1,19)'	8	0
'u(k-4,4)*u(k-5,6)'	9	0
'u(k-3,11)*u(k-4,19)'	10	0
	11	5
	12	10
	13	5
	14	8
	15	0
	16	0
	17	3
	18	0
	19	5
	20	0

Table B.18: Production Well 23

Production well 24	Injection Well	Net Ranking
'y(k-1)'	1	6
'u(k-2,13)*y(k-1)'	2	0
'u(k-3,4)*u(k-5,4)'	3	0
'u(k-1,18)*u(k-2,19)'	4	1
'u(k-3,11)*y(k-1)'	5	0
'u(k-1,18)*u(k-5,19)'	6	0
'u(k-4,18)*u(k-2,19)'	7	0
'u(k-2,1)*u(k-4,4)'	8	0
	9	0
	10	0
	11	5
	12	0
	13	4
	14	0
	15	0
	16	0
	17	0
	18	2
	19	2
	20	0

Table B.19: Production Well 24

Appendix C

Future Injection Values

The following tables show the estimated injection values for increasing future oil production in one year period. Injection rates are given row-wise for the corresponding well (20 injection wells, 12 future months).

Injection Well Month	1	2	3	4	5	6	7	8	9	10	11	12
1	0.5969	0.0851	0.9738	0.2321	0.2138	0.1572	0.9156	0.6291	0.4617	0.4350	0.1334	0.7132
2	0.7580	0.5572	0.4000	0.2258	0.2079	0.9890	0.8538	0.4432	0.4851	0.8473	0.4211	0.9911
3	0.8316	0.9454	0.8741	0.8793	0.5354	0.8444	0.0373	0.7835	0.9773	0.9944	0.2405	0.9455
4	0.9834	0.1043	0.3703	0.6628	0.6054	0.8305	0.1172	0.6406	0.0424	0.2625	0.7296	0.8499
5	0.7239	0.1200	0.5960	0.2916	0.6852	0.0657	0.0962	0.4944	0.3737	0.9439	0.4626	0.7902
7	0.3767	0.9962	0.9236	0.2074	0.6092	0.0899	0.9490	0.5139	0.9221	0.8352	0.6833	0.8005
13	0.9803	0.8479	0.0878	0.6373	0.0500	0.1132	0.2386	0.0445	0.6087	0.1316	0.5712	0.0272
14	0.3482	0.7599	0.3721	0.8402	0.3189	0.0154	0.1502	0.6323	0.6826	0.1787	0.0796	0.5540
15	0.3315	0.1890	0.9975	0.6968	0.0373	0.6285	0.6111	0.6672	0.0441	0.4440	0.2299	0.7653
16	0.7865	0.5377	0.2070	0.7243	0.8493	0.6327	0.9198	0.8412	0.4344	0.9702	0.5983	0.2107
18	0.0211	0.0164	0.0077	0.0863	0.0074	0.0441	0.0678	0.0703	0.1856	0.0198	0.0094	0.0208
19	0.1169	0.8629	0.1123	0.3556	0.2394	0.4977	0.1673	0.4776	0.8038	0.1781	0.4724	0.5304

Table C.1: Injection Rates for Max Oil Production-Polynomial Model

Injection Well Month	1	2	3	4	5	6	7	8	9	10	11	12
1	0.0817	0.7069	0.5170	0.0351	0.0388	0.0771	0.8148	0.1468	0.6980	0.9790	0.3909	0.0218
2	0.7591	0.4332	0.8886	0.6558	0.2019	0.3273	0.7802	0.4946	0.6814	0.9484	0.4660	0.7616
3	0.9738	0.7877	0.7633	0.3531	0.5189	0.3401	0.4805	0.2125	0.8662	0.4711	0.7902	0.8318
4	0.8948	0.5170	0.6996	0.6021	0.0235	0.7880	0.8456	0.2636	0.4294	0.1766	0.9173	0.4385
5	0.6448	0.9016	0.4665	0.1973	0.7069	0.2049	0.4818	0.5907	0.4211	0.7728	0.9160	0.4288
7	0.3010	0.6535	0.9222	0.3708	0.8280	0.1206	0.8820	0.6263	0.6985	0.1643	0.1475	0.3211
13	0.7122	0.6093	0.7538	0.5731	0.2269	0.3373	0.9907	0.3073	0.0949	0.6978	0.3263	0.5158
14	0.6270	0.8034	0.6376	0.6364	0.2388	0.8591	0.8126	0.0526	0.3564	0.0775	0.0516	0.1134
15	0.2074	0.0090	0.8399	0.4387	0.3353	0.9158	0.8158	0.2892	0.5982	0.8090	0.7804	0.1059
16	0.8006	0.9438	0.0227	0.0123	0.9578	0.0255	0.0567	0.1122	0.7191	0.9790	0.5334	0.6793
17	0.6401	0.1652	0.6648	0.4839	0.6685	0.9849	0.4106	0.2921	0.7402	0.4061	0.5381	0.6055
18	0.1523	0.0261	0.1022	0.3360	0.0599	0.4475	0.1307	0.0722	0.0089	0.1565	0.3546	0.5095
19	0.7610	0.3032	0.5949	0.4852	0.8050	0.6251	0.9308	0.6490	0.2681	0.9411	0.9448	0.0650

Table C.2: Injection Rates for 30% Increase in Oil Production-Polynomial Model

Injection Well Month	1	2	3	4	5	6	7	8	9	10	11	12
1	0.0764	0.0041	0.0116	0.9982	0.9548	0.9203	0.9724	0.9644	0.0447	0.0191	0.9862	0.9908
2	0.2595	0.1846	0.4961	0.4443	0.6165	0.4148	0.2457	0.7592	0.5150	0.4602	0.5548	0.4546
3	0.7068	0.3548	0.7629	0.4819	0.1881	0.5514	0.4509	0.7856	0.9084	0.8451	0.6744	0.2654
4	0.0879	0.3839	0.1679	0.1969	0.0183	0.1277	0.0901	0.1014	0.1002	0.0729	0.2508	0.1325
5	0.2476	0.0800	0.1095	0.4624	0.4186	0.1453	0.2909	0.2586	0.3355	0.0952	0.2463	0.0033
6	0.9357	0.2121	0.6544	0.1052	0.7745	0.3509	0.1294	0.0519	0.1340	0.0808	0.3707	0.1845
7	0.4337	0.6634	0.5130	0.4413	0.7031	0.3044	0.6229	0.7241	0.6553	0.5544	0.5283	0.5375
8	0.6311	0.3235	0.4438	0.3197	0.9615	0.5987	0.7408	0.3890	0.5184	0.4331	0.5721	0.1121
9	0.2153	0.5780	0.1964	0.0872	0.2349	0.1945	0.1859	0.2598	0.2210	0.0283	0.0278	0.1664
10	0.4152	0.0210	0.2686	0.0383	0.0898	0.1297	0.2440	0.3820	0.2741	0.2329	0.1090	0.0218
11	0.3272	0.2354	0.4441	0.3657	0.6776	0.1394	0.2395	0.0332	0.0837	0.4798	0.0720	0.2553
12	0.7310	0.4433	0.6484	0.8042	0.7664	0.6806	0.6964	0.9190	0.5265	0.8577	0.4222	0.5118
13	0.0225	0.2529	0.3013	0.1440	0.1390	0.3341	0.4950	0.6774	0.7998	0.4419	0.4982	0.5625
14	0.4808	0.5426	0.6688	0.7004	0.6213	0.4372	0.5431	0.6345	0.5526	0.8662	0.5063	0.2806
15	0.5742	0.1524	0.3190	0.3832	0.6712	0.7042	0.3842	0.0876	0.4434	0.3620	0.2893	0.5766
16	0.3303	0.9906	0.8420	0.6157	0.7506	0.5162	0.7101	0.7580	0.5781	0.6311	0.6892	0.8187
17	0.2339	0.8869	0.2123	0.1551	0.1894	0.0080	0.5373	0.3515	0.3505	0.3949	0.4126	0.1682
18	0.2089	0.3109	0.1614	0.0535	0.6386	0.0758	0.1245	0.0047	0.0241	0.2937	0.3559	0.0277
19	0.3169	0.1646	0.0507	0.4327	0.2928	0.1890	0.1945	0.4187	0.1706	0.2372	0.1138	0.2386
20	0.2229	0.2502	0.1111	0.0323	0.1507	0.3859	0.3315	0.1416	0.2335	0.0461	0.3339	0.4396

Table C.3: Injection Rates for Max Oil Production-MSRBF Model

Injection Well Month	1	2	3	4	5	6	7	8	9	10	11	12
1	0.4157	0.0213	0.9423	0.9374	0.0618	0.9312	0.9892	0.8911	0.9290	0.9558	0.9349	0.9528
2	0.2184	0.4238	0.0616	0.6427	0.1887	0.3780	0.5486	0.3458	0.3029	0.9261	0.9392	0.4308
3	0.5988	0.7263	0.5150	0.3747	0.5080	0.8484	0.4481	0.7794	0.4213	0.4794	0.1018	0.0093
4	0.0940	0.4170	0.2852	0.2050	0.1503	0.5965	0.2607	0.0230	0.1022	0.4252	0.1991	0.1361
5	0.1153	0.1014	0.1793	0.9970	0.5598	0.2933	0.0417	0.1513	0.1220	0.3274	0.3051	0.0289
6	0.5244	0.1951	0.6696	0.1440	0.2915	0.5222	0.8025	0.5822	0.5469	0.0791	0.4536	0.6180
7	0.8058	0.9238	0.5869	0.5893	0.8058	0.4850	0.7282	0.3201	0.4727	0.4945	0.5845	0.7007
8	0.6914	0.1888	0.6328	0.2961	0.1082	0.4848	0.2239	0.6501	0.3755	0.7367	0.7060	0.9281
9	0.7178	0.1075	0.1765	0.3113	0.6290	0.2836	0.1170	0.2493	0.1191	0.6409	0.4156	0.6392
10	0.5355	0.3790	0.5169	0.1008	0.0189	0.2894	0.4708	0.8824	0.6432	0.6260	0.3257	0.4095
11	0.2454	0.2487	0.5259	0.1237	0.0107	0.4153	0.1793	0.1020	0.1262	0.2986	0.4020	0.1877
12	0.3444	0.0596	0.8745	0.7461	0.6397	0.3686	0.9667	0.4521	0.7596	0.7148	0.3821	0.6850
13	0.1350	0.2837	0.1642	0.1472	0.0032	0.4896	0.4206	0.4716	0.3182	0.7684	0.4712	0.6200
14	0.7143	0.8448	0.9123	0.5472	0.3778	0.5084	0.2798	0.8249	0.7583	0.6477	0.9111	0.4805
15	0.0538	0.2863	0.6558	0.2768	0.4089	0.7722	0.3841	0.7733	0.7762	0.1818	0.0853	0.6827
16	0.6662	0.5654	0.4356	0.7919	0.8314	0.8736	0.2936	0.9868	0.5535	0.8635	0.7455	0.8797
17	0.3297	0.8006	0.5621	0.3809	0.3990	0.3307	0.4053	0.4012	0.6297	0.5895	0.2116	0.2065
18	0.2623	0.2082	0.3568	0.5956	0.4018	0.2931	0.6209	0.1046	0.2003	0.2348	0.2250	0.2881
19	0.0077	0.2812	0.2020	0.2019	0.2263	0.1083	0.2650	0.0304	0.6953	0.2027	0.4254	0.1039
20	0.2085	0.0765	0.2253	0.3043	0.2891	0.5816	0.2630	0.1143	0.2803	0.0100	0.3658	0.0462

Table C.4: Injection Rates for 30% Increase Oil Production-MSRBF Model

Bibliography

[1] British Petroleum, "BP Statistical Review of World Energy, June 2014," London, Tech. Rep. June, 2014, accessed 25/07/2016. [Online]. Available: http://www.bp.com/en/global/corporate/about-bp/energy-economics/statistical-review-of-world-energy.html

[2] WEC, "World Energy in 4 Minutes," 2016, accessed 2017-01-02. [Online]. Available: https://www.worldenergy.org/

[3] U.S. Energy Information Administration, "Annual Energy Outlook 2015," Tech. Rep., 2015.

[4] C. O. F. Technology, "Three Phases of Oil Recovery," 2007, accessed 10/08/205. [Online]. Available: http://www.chinaoilfieldtech.com/oilrecovery.html

[5] E. L. Anadón, V. Casalotti, and G. Masarik, "El abecé de los Hidrocarburos No Convencionales," Instituto Argentino del Petróleo y Gas, Buenos Aires, Tech. Rep., 2013.

[6] V. Alvarado and E. Manrique, "Enhanced oil recovery: An update review," *Energies*, vol. 3, no. 9, pp. 1529–1575, 2010.

[7] TOTAL. S.A, "Mature Fields Inventing the Future," mar 2007.

[8] K. Holdaway, *Harness Oil and Gas Big Data with Analytics*, 1st ed. New Jersey: John Wiley & Sons, Ltd, 2014.

[9] Comision Nacional de Hidrocarburos, "DOCUMENTO TECNICO 2 (DT2) La Tecnologia de Exploracion y Produccion en Mexico y en el Mundo: Situacion Actual y Retos," Comision Nacional de Hidrocarburos, Mexico D.F, Tech. Rep., 2011.

[10] R. Gaskari, S. Mohaghegh, and J. Jalali, "An Integrated Technique for PDA with Application to Mature Fields," *SPE Production & Operations*, vol. 22, no. 4, pp. 15–18, 2007.

[11] M. Arianson, "Well Models for Production Optimization," Master Thesis, Norwegian University of Science and Technology, 2008.

[12] L. V. Branets, S. S. Ghai, S. L. Lyons, and X. H. Wu, "Challenges and technologies in reservoir modeling," *Communications in Computational Physics*, vol. 6, no. 1, pp. 1–23, 2009.

[13] D. Kaviani, "Interwell Connectivity Evaluation From a Wellrate Fluctuations: A Waterflooding Managment Tool," PhD Thesis, Texas A&M, 2009.

[14] R. Belkis T and L. White, "Reservoir Characterization Based on Tracer Response and Rank Analysis of Production and Injection Rates," in *The American Association of Petroleum Geologists*, 1999, pp. 209–218.

[15] S. K. Matthai, "Reservoir Simulation: Mathematical Techniques in Oil Recovery," *Geofluids*, vol. 8, no. 4, pp. 344–345, 2008.

[16] M. Höök, "Depletion and decline curve analysis in crude oil production," Ph.D. dissertation, Uppsala University, 2009.

[17] L. P. Dake, *Fundamentals of Reservoir Engineering*, 17th ed. The Hague: Elsevier, 1998.

[18] J. J. Arps, "Analysis of Decline Curves," *Transactions of the AIME*, vol. 160, no. 1, pp. 228–247, 1945, accessed 03/03/2016. [Online]. Available: http://www.onepetro.org/doi/10.2118/945228-G{\%}5Cnhttp://dx.doi.org/10.2118/945228-G

[19] K. Holdaway, "Let Oil and Gas Talk to you: Predicting Production Performance," in *SAS Global Forum 2012*, Cary, 2012, pp. 1–12.

[20] G. L. Chierici, "The Material Balance Equation," in *Principles of Petroleum Reservoir Engineering*, Berlin, Heidelberg, 1994, pp. 359–404.

[21] J. Kleppe, "TPG4150 Reservoir Recovery Techniques 2016 Material Balance Equations," Norwegian University of Science and Technology, Oslo, Tech. Rep., 2016.

[22] L. P. Dake, *Fundamentals of Reservoir Engineering*, 1st ed. The Hague: Elsevier, 1978.

[23] A. A. Al-Yousef, "Investigating Statistical Techniques To Infer Interwell Connectivity From Production and Injection Rate Fluctuations," PhD Thesis, The University of Texas, Austin, 2006.

[24] R. Laochamroonvorap, "Advances in the development and application of a capacitance-resistance model," Master Thesis, The University of Texas at Austin, 2013, accessed 03/08/2016. [Online]. Available: http://repositories.lib.utexas.edu/handle/2152/22372

[25] G. a. Moreno, "Multilayer capacitance resistance model with dynamic connectivities," *Journal of Petroleum Science and Engineering*, vol. 109, pp. 298–307, 2013, accessed 02/08/2016. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0920410513001848

[26] A. Al-Yousef and D. Weber, "Interwell Connectivity and Diagnosis Using Correlation of Production and Injection Rate Data in Hydrocarbon Production," Texas A&M University, College Station, Tech. Rep. March, 2007.

[27] I. G. Main, L. Li, K. J. Heffer, O. Papasouliotis, and T. Leonard, "Long-range, critical-point dynamics in oil field flow rate data," *Geophysical Research Letters*, vol. 33, no. 18, pp. 1–5, 2006.

[28] I. G. Main, L. Li, K. J. Heffer, O. Papasouliotis, T. Leonard, N. C. Koutsabeloulis, and X. Zhang, "The Statistical Reservoir Model: calibrating faults and fractures, and predicting reservoir response to water flood," *Geological Society, London, Special Publications*, vol. 292, no. 1, pp. 469–482, 2007.

[29] F. Hoffmann, D. J. Hand, N. Adams, D. Fisher, and G. Guimaraes, Eds., *Correlation-Based and Contextual Merit-Based Ensemble Feature Selection*. Berlin, Heidelberg: Springer, 2001, accessed 17/05/2017. [Online]. Available: http://dx.doi.org/10.1007/3-540-44816-0{_}14

[30] I. Main, D. Richardson, and L. Ormerod, "THE UNIVERSITY OF EDINBURGH STATISTICAL RESERVOIR ANALYSIS WHITEPAPER," The University of Edinburgh, Edinburgh, Tech. Rep. May, 2011.

[31] C. Chatfield, *The Analysis of Time Series: an Introduction*, sixth edit ed. London: Chapman & Hall, 2004, vol. 140, no. 2.

[32] E. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*, 4th ed. Chicago: Willey-Blackwell, 2008, vol. 4, accessed 06/04/2017. [Online]. Available: http://doi.wiley.com/10.1111/j.1467-9892.2009.00643.x

[33] L. Ljung, *System Identification: Theory for User*, 1987, vol. 11, no. 3.

[34] G. Van Essen, P. Van den Hof, and J.-D. Jansen, "A two-Level Strategy to Realize Life-Cycle Production Optimization in an Operational Setting," *Society of Petroleum Engineers*, vol. 18, no. 6, pp. 1057–1066, 2013.

[35] L. Kun-Han, "Investigating Statistical Modelling Approaches for Reservoir Characterisation in Waterfloods from Rates Fluctuations," PhD Thesis, University of Southern California, 2010.

[36] G. DeJonge and M. Stundner, "How to Routine Reservoir Surveillance with Neural Networks and Simplified Reservoir Models Can Convert Data Into Intormation," in *European Petroleum Conference*. Aberdeen: SPE, 2002.

[37] L. Saputell, H. Malki, J. Canelon, and N. M, "A Critical Overview of Arti cial Neural Network Applications in the Context of Continuous Oil Field Optimization." in *SPE Annual Technical Conference and Exhibition*. San Antonio,TX: SPE, 2002.

[38] S. Esmaili and S. D. Mohaghegh, "Full field reservoir modeling of shale assets using advanced data-driven analytics," *Geoscience Frontiers*, vol. 7, no. 1, pp. 11–20, 2016.

[39] A. Tahar, "Neural network applications to reservoirs: Physics-based models and data models," *Journal of Petroleum Science and Engineering*, pp. 1–6, 2014.

[40] T. Poggio and G. Federico, "Networks for approximation and learning," *Proceedings of the IEEE*, vol. 78, no. 9, 1990.

[41] I. Croall and J. Mason, "Industrial Applications of Neural Networks," AEA Technology, Didcot, Tech. Rep., 1992.

[42] M. T. Hagan and M. B. Menhaj, "No Title," *Transactions on Neural Networks*, vol. 5, no. 6, 1994.

[43] J. J. More, "The Levenberg-Marquardt algorithm: Implementation and theory," *Lecture Notes in Mathematics*, vol. 630, pp. 105–116, 1978.

[44] J. F. Khaw, B. Lim, and L. E. Lim, "Optimal design of neural networks using the Taguchi method," *Neurocomputing*, vol. 7, no. 3, pp. 225–245, 1995, accessed 05/13/2016. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/092523129400013I

[45] M. Bashiri and A. Farshbaf Geranmayeh, "Tuning the parameters of an artificial neural network using central composite design and genetic algorithm," *Scientia Iranica*, vol. 18, no. 6, pp. 1600–1608, 2011, accessed 16/05/2017. [Online]. Available: http://dx.doi.org/10.1016/j.scient.2011.08.031

[46] A. J. Maren, C. T. Harston, R. M. Pap, A. J. Maren, D. Jones, and S. Franklin, "CONFIGURING AND OPTIMIZING THE BACK-PROPAGATION NETWORK," in *Handbook of Neural Computing Applications*, 1st ed. San Diego, CA: Academic Press Inc, 1990, pp. 233–250.

[47] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015, accessed 03/03/2017. [Online]. Available: http://dx.doi.org/10.1016/j.neunet.2014.09.003

[48] Q. V.L, M. Ranzato, R. Monja, M. Devin, K. Chen, G. Corrado, J. Dean, and N. Andrew, "Building High-level Features Using Large Scale Unsupervised Learning," in *International Conference on Machine Learning*, Edinburgh, 2012.

[49] Derrick, "Derrick A leading provider of business research for the Oil and Gas Industry," Derrick, Tech. Rep., 2013, accessed 17/05/2017. [Online]. Available: http://www.1derrick.com/reports/{\#}

[50] L. Fortuna, S. Graziani, A. Rizzo, and M. G. Xibilia, *Soft Sensors for Monitoring and Control of Industrial Processes*, 1st ed. London: Springer, 2007.

[51] J. W. Graham, "Missing data analysis: making it work in the real world." *Annual review of psychology*, vol. 60, pp. 549–76, 2009, accessed 03/03/2017. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/18652544

[52] F. Guo, "A New Identification Method for Wiener and Hammerstein Systems," PhD Thesis, Universität Karlsruhe, 2003. [Online]. Available: www.ubka.uni-karlsruhe.de/volltexte/fzk/6955/6955.pdf?

[53] B. Mirkin, *Core Concepts in Data Analysis: Summarization, Correlation and Visualization*, 1st ed. London: Springer-Verlag London, 2011.

[54] S. Billings and S. Chen, "The determination of of multivariable nonlinear models for dynamic systems using neural networks," *Neural Network Systems Techniques and Applications*, 1998.

[55] S. Billings, *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*, 1st ed. Singapore: John Wiley & Sons, Ltd, 2013.

[56] S. Billings and H. L. Wei, "Sparse model identification using a forward orthogonal regression algorithm aided by mutual information," *IEEE Transactions on Neural Networks*, vol. 18, no. 1, pp. 306–310, 2007.

[57] U. Nehmzow, O. Akanyeti, and S. A. Billings, "Towards modelling complex robot training tasks through system identification," *Journal of Robotics and Autonomous Systems*, vol. 58, pp. 265–275, 2010.

[58] M. Balikhin, R. Boynton, S. Walker, J. Borovsky, S. A. Billings, and H. Wei, "Using the NARMAX approach to model the evolution of energic electrons fluxes at geostationary orbit," *Geophisical Research Letters*, vol. 38, p. L18105, 2011.

[59] Y. Zhao, S. A. Billings, and H. Wei, "A new fast cellular automata neighbourhood detection and rule identification algorithm," *IEEE*

Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 4, pp. 1283–1287, 2012.

[60] E. Vidal-Rosas, "Advanced tomographic reconstruction methods for diffuse optical imaging of cerebral heamodynamic response," PhD Thesis, Unversity of Sheffield, 2011.

[61] U. Friederich, S. A. Billings, R. Hardie, J. M, and D. Coca, "Fly Photoreceptors Encode Phase Congruency," *PLOS One*, vol. 11, no. 6, 2016.

[62] K. Krishnanathan, S. Anderson, and S. A. Billings, "A data-driven framework for identifying nonlinear dynamic models of generic parts," *ACS Synthetic Biology*, vol. 8, pp. 375–384, 2012.

[63] H. Wei and S. A. Billings, "An efficient nonlinear cardinal B-spine model for high tide forecasts at the Venice Lagoon," *Nonlinear Processes in Geophysics*, vol. 13, pp. 577–584, 2006.

[64] X. Guyon and J.-f. Yao, "On the Underfitting and Overfitting Sets of Models Chosen by Order Selection Criteria," *Journal of Multivariate Analysis*, vol. 70, pp. 221–249, 1999.

[65] C. M. Bishop, "Model-based machine learning." *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 371, no. 1984, pp. 2012–2022, 2013.

[66] S. a. Billings, S. Chen, and M. J. Korenberg, "Identification of MIMO non-linear systems using a forward-regression orthogonal estimator," *Int. J. Control*, vol. 49, no. 6, pp. 2157–2189, 1989, accessed 16/05/2017. [Online]. Available: http://eprints.soton.ac.uk/251146/

[67] S. Chen, S. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of Control*, vol. 50, no. 5, pp. 1873–1896, 1989. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/00207178908953472

[68] H. L. Wei, S. A. Billings, and J. Liu, "Term and variable selection for non-linear system identification," *International Journal of Control*, vol. 77, no. 1, pp. 86–110, 2004.

[69] M. Sayarpour, E. Zuluaga, C. S. Kabir, and L. W. Lake, "The use of capacitance-resistance models for rapid estimation of waterflood performance and optimization," *Journal of Petroleum Science and Engineering*, vol. 69, no. 3-4, pp. 227–238, 2009, accessed 01/23/2017. [Online]. Available: http://dx.doi.org/10.1016/j.petrol.2009.09.006

[70] S. Li and X. Liang, "A state space model to infer interwell connectivity only form injection and production data in waterfloods," *2009 2nd International Conference on Intelligent Computing Technology and Automation, ICICTA 2009*, vol. 1, no. 1, pp. 949–952, 2009.

[71] Y. Ali, J. L. Jensen, and L. W. Lake, "Integrated interpretation of interwell connectivity using injection and production fluctuations," *Mathematical Geosciences*, vol. 41, no. 1, pp. 81–102, 2009.

[72] X. Liang, "A simple model to infer interwell connectivity only from well-rate fluctuations in waterfloods," *Journal of Petroleum Science and Engineering*, vol. 70, no. 1-2, pp. 35–43, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.petrol.2009.08.016

[73] T. G. Neiting, "Making and Evaluating Point Forecasts," *Journal of the American Statistical Association*, vol. 106, no. 494, pp. 746–762, 2011.

[74] Nexen, "Other UK Operations," 2016, accessed 04/02/2016. [Online]. Available: http://www.nexencnoocltd.com/en/Operations/Conventional/UKNorthSea///OtherUKOperations.aspx

[75] J. Park and I. Sandberg, "Universal Approximation using Radial-Basis-Function Networks," *Neural Computation*, vol. 3, no. 2, pp. 246–257, 1991.

[76] T. Kohonen, "An introduction to neural computing," *Neural Networks*, vol. 1, no. 1, pp. 3–16, 1988.

[77] S. Chen, S. a. Billings, C. F. N. Cowan, and P. M. Grant, "Practical identification of NARMAX models using radial basis functions," *International Journal of Control*, vol. 52, no. 769892610, pp. 1327–1350, 1990.

[78] S. Chen and S. Billings, "Neural networks for nonlinear dynamic system modelling and identification," *International Journal of Control*, vol. 56, no. 2, pp. 319–346, 1992.

[79] S. Billings, H.-L. Wei, and M. A. Balikhin, "Multiscale Radial Basis Function Networks," Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, Tech. Rep., 2005.

[80] H. L. Wei and S. Billings, "Feature subset selection and ranking for data dimensionality reduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 162–166, 2007.

[81] S. McLoone, M. Brown, G. Irwin, and L. A, "A hybrid linear/nonlinear training algorithm for feedforward neural networks," *IEEE Transactions on Neural Networks*, vol. 9, no. 4, pp. 669–684, 1988.

[82] S. Chen, X. Hong, C. J. Harris, and P. M. Sharkey, "Sparse Modeling Using Orthogonal Forward Regression With PRESS Statistic and Regularization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 2, pp. 898–911, 2004.

[83] G. Zheng and S. Billings, "Radial Basis Network Configuration Using Mutual Information and Orthogonal Least Squares," *Journal of Neural Networks*, vol. 9, pp. 1619–1637, 1996.

[84] S. Chen, S. Billings, and P. Grant, "Recursive hybrid algorithm for non-linear system identification using radial basis function networks," *International Journal of Control*, vol. 55, pp. 1051–1070, 1992.

[85] S. Chen, S. Billings, C. Cowan, and P. Grant, "Practical Identification of NARMAX Models Using Radial Basis Functions," The University of Sheffield, Sheffield, Tech. Rep., 1990.

[86] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout : A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research (JMLR)*, vol. 15, pp. 1929–1958, 2014.

[87] M. Gethsiyal Augasta and T. Kathirvalavakumar, "Pruning algorithms of neural networks- a comparative study," *Central European Journal of Computer Science*, vol. 3, no. 3, pp. 105–115, 2013.

[88] D. Arnold, V. Demyanov, D. Tatum, M. Christie, T. Rojas, S. Geiger, and P. Corbett, "Hierarchical benchmark case study for history matching , uncertainty quantification and reservoir characterisation," *Computers and Geosciences*, vol. 50, pp. 4–15, 2013, accessed 15/03/2017. [Online]. Available: http://dx.doi.org/10.1016/j.cageo.2012.09.011

[89] Y. Hajizadeh, "Population-Based Algorithms for Improved History Matching and Uncertainty Quantification of Petroleum Reservoirs," Ph.D. dissertation, Heriot Watt University, 2011.

[90] D. Blockley and P. Godfrey, *Doing it Differently: Systems for Rethinking Construction*, 1st ed. Thomas Telford, 2000.

[91] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *Journal of Machine Learning Research (JMLR)*, vol. 3, no. 3, pp. 1157–1182, 2003.

[92] A. Saltelli, M. Ratto, T. Amdres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis. The Primer*, 1st ed. Sussex: John Wiley & Sons, Ltd, 2008.

[93] M. Kudo and J. Sklansky, "Comparison of algorithms that select features for pattern classifiers," *Pattern Recognition*, vol. 33, no. 1, pp. 25–41, 2000.

[94] M. Kuhn and K. Johnson, *Applied Predictive Modeling*, 1st ed. Springer, 2013, accessed 15/03/2017. [Online]. Available: http://link.springer.com/10.1007/978-1-4614-6849-3

[95] M. S. Joshi, *The Concepts and Practice of Mathematical Finance*, 2nd ed. Cambridge University Press, 2008.

[96] K. Heldman, *Risk Management: Project Managers Spotlight on*, 1st ed., L. E. Light, Ed. Alameda: Neil Edde, 2005.

[97] R. C. Grinold and R. N. Kahn, *Active Portfolio Management : A Quantitative Approach for Providing Superior Returns and Controlling Risk*, 2nd ed., ser. [Irwin Library of Investment & Finance]. New York: McGraw-Hill, 1999, accession Number: 52240; Language: English.

[98] B. Möller and U. Reuter, *Uncertanty Forecasting in Engineering*, 1st ed. Dresden: Springer, 2007.

[99] C. de Madrid, "Risk Analysis and Quantification," Comunidad de Madrid, Madrid, Tech. Rep., 2012.

[100] W. H. Organization, "Part 1 : Guidance Document on Characterizing and Communicating Uncertainty in Exposure Assessment Part 2 : Hallmarks of Data Quality in Chemical," *WHO*, 2008.

[101] M. Zoveidavianpoor and M. Jalilavi, "Qualitative Analysis of Enhanced Oil Recovery : Impacts on Air , Surface Water and Groundwater," in *International Conference on Chemical, Environment & Biological Sciences*, Kuala Lumpur, 2014, pp. 11–15, accessed 17/05/2017. [Online]. Available: http://iicbe.org/upload/8785C914075.pdf

[102] K. Ishikawa, *Introduction to Quality Control*, 1st ed. London: Chapman & Hall, 1990.

[103] A. M. Daniel, S. Jay, J. K. Myoung, S. Dale, and Thonhteeraparp, "Number of Replications Required in Monte Carlo Simulation Studies:

A Synthesis of Four Studies," *Journal of Modern Applied Statistical Methods*, vol. 10, no. 1, 2011.

[104] A. Slaughter, "Developing oil and natural gas supply models to meet user needs," Energy Information Administration, Tech. Rep., 2014.

[105] Oxford University Press, "Oxford Dictionary Online," 2014, accessed 03/03/2016. [Online]. Available: http://www.oxforddictionaries.com/

[106] M. Asadollahi, "Waterflooding Optimization for Improved Reservoir Management Masoud Asadollahi," PhD Thesis, Norwegian University of Science and Technology, 2012.

[107] H. Agus, "Optimization and Control of Petroleum Reservoirs," PhD Thesis, Norwegian University of Science and Technology, Trondheim, 2013.

[108] G. Venter, "Review of Optimization Techniques," Stellencosch University, Matieland, Tech. Rep., 2010.

[109] M. Gen and C. Ribeiro, *Handbook of Metaheuristics*, 2nd ed., M. Gendreau and J.-Y. Potvin, Eds. Springer, 2011, vol. 157.

[110] M. López-Ibáñez and T. Stützle, "Automatically improving the anytime behaviour of optimisation algorithms," *European Journal of Operational Research*, vol. 235, no. 3, pp. 569–582, 2014.

[111] T. A. Feo and M. G. C. Resende, "Greedy Randomized Adaptive Search Procedures," *Journal of Global Optimization*, vol. 6, no. 2, pp. 109–133, 1995.

[112] P. Festa and M. G. C. Resende, "An annotated bibliography of GRASP-Part II: Applications," *International Transactions in Operational Research*, vol. 16, no. 2, pp. 131–172, 2009, accessed 03/12/2016. [Online]. Available: http://doi.wiley.com/10.1111/j.1475-3995.2009.00664.x

[113] S. D. Rahmawati, "Integrated Field Modeling and Optimization," Ph.D. dissertation, Norwegian University of Science and Technology, 2012.

[114] G. C. Silva, L. S. Ochi, and S. L. Martins, "Experimental comparison of greedy randomized adaptive search procedures for the maximum diversity problem," *Proceedings of the 3rd international workshop on efficient and experimental algorithms (WEA 2004), Lectures Notes on Computer Science (LNCS)*, vol. 3059, pp. 498–512, 2004, accessed 10/12/2016. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-24944501526{\&}partnerID=tZOtx3y1

[115] R. Braal, "Production Optimization," TNO, Delf, Tech. Rep., 2012.

[116] Nastaq, "Crude Oil Brent," 2016, accessed 2016-08-03. [Online]. Available: http://www.nasdaq.com/markets/crude-oil-brent-b.aspx