UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

A Comparative Study between Proposed Hybrid and known Decline Curve Models and Financial Impacts

PhD RESEARCH THESIS

Prepared by

Prinisha Manda

690765

Submitted to

School of Chemical and Metallurgical Engineering, Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, South Africa

Supervisor(s): Prof Diakanua Nkazi

December 2021

Candidate's Declaration

I declare that this thesis is my own unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Prinisha Manda (690765)

13th day of October in the year 2021

Abstract

The development of prediction tools for production performance and the lifespan of shale gas reservoirs has been a focus for petroleum engineers. Several decline curve models have been developed and compared with data from shale gas production. Initially in this study, the current or existing decline curve models were evaluated using the goodness of fit as a measure of accuracy with field data.

The evaluation found that there are advantages in using the current decline curve models; however, they also have limitations associated with them, which have to be addressed. A new hybrid model, which incorporates the Autoregressive Integrated Moving Average (ARIMA), and the Artificial Neural Network (ANN), was examined and reasoned from literature to provide a higher level of accuracy. Based on the accuracy assessment conducted on the different models, the Stretched Exponential Decline Model (SEDM) and Logistic Growth Model (LGM), followed by the Exponential Decline Model (EDM), the Power Law Exponential Model (PLE), the Duong Model and, lastly, the Arps Hyperbolic Decline Model provide the best fit with production data. The coefficient of variance (R^2) values were, 0.9672, 0.9627, 0.9528, 0.9512, 0.9382 and 0.8849, respectively.

Secondly, the study used the hybrid model philosophy to develop, predict and validate shale gas decline. The results indicated that the PLE and Duong model provided the best-fit and R^2 with the estimated data. The use of hybrid models provides a more precise predicting model for forecasting time series data, as compared to an individual model.

The forecasting performance of decline curve hybrid models and ANN-ARIMA hybrid models are evaluated and compared with Arps, Duong, PLE, ARIMA and ANN models, respectively. The variable used to assess the models was the respective flow rate, q(t) monitored over a period of time (t). The results have shown that the Arps-PLE hybrid decline model had the lowest root mean square (RMSE) and good R^2 followed by the ANN and ARIMA models. The result provided a significant contribution to the prediction of shale gas production. The Arps-PLE hybrid decline model is a good model predictor for shale gas production. The contributing factor is the dominance of the PLE parameters i.e., D_i changes at early stages and D_∞ become constant at late

time in the model. This caters for the transient flow regime (TFR) which the Arps decline model did not consider.

Thirdly and lastly, the study evaluated the EUR, and it was found that different values are obtained from the various models. The EUR is either over or underestimated. The Arps-PLE hybrid decline and ANN models, which were found to be the best models in predicting values closest to the actuals, were used to calculate the EUR and to compare with other decline curve models. The results clearly show the overestimation of the EUR values for the different shale plays using the Arps, Duong and PLE decline models, compared to the Arps-PLE hybrid decline and ANN models. Evaluating the EUR accurately would then allow for the accurate estimation of the total net revenue generated from a shale play.

Dedication

In loving memory of my dad, the late Harry Moodley

Forever in my heart …

"Strength does not come from winning; your struggles develop your strengths. When you go through hardships and decide not to surrender, that is strength." – Mahatma Gandhi

Publications and Presentations

Publications[1]

1. The Evaluation and Sensitivity of Decline Curve Modelling, Prinisha Manda and Diakanua Nkazi, *Energies (MDPI)*, *13*(11), p 2765.
2. Production Decline Prediction of Shale Gas using Hybrid Models, Prinisha Manda and Diakanua Nkazi, *Global Journal of Researches in Engineering (GJRE)*, *20*(5), pp. 1-15.
3. Assessing the Accuracy and Validity of Single and Hybrid Models for Predicting Shale Gas Production – Prinisha Manda and Diakanua Nkazi, *Energy and Fuels*, *35*(7), pp. 6068-6080.
4. Financial Modelling of Shale Gas, Prinisha Manda and Diakanua Nkazi, In-Progress.

Presentations

1. The Evaluation and Sensitivity of Decline Curve Modelling, Prinisha Manda and Diakanua Nkazi, 70th Canadian Chemical Engineering Conference (CCEC) (30 October 2020).

[1] Publications can be found in Appendix B

Acknowledgments

Firstly, and most importantly, I would like to honour my lord and saviour Jesus Christ, without him, nothing would be possible. I reference, **Philippians 4:13** and I quote: "I can do all this through him who gives me strength."

I wish to thank Professor Diakanua Nkazi for agreeing to be my supervisor.

I would like to thank KAPPA and JMP for allowing me access to the use of their software for this study.

Dad, you gave me the greatest gift anyone could give another person, you believed in me. Your unconditional love moulded me into a strong and independent woman. I will always be thankful and grateful to you. I love and miss you.

Mum, I am a motivated and resilient woman because you raised me to conquer anything life puts in front of me. I honour, love and cherish you dearly.

Justin and Liam, "My boys", you both motivate and inspire me every day to be the best wife and mother. This journey would not be a reality without your love and support. I treasure and love you both immensely.

Table of Contents

List of Figures

List of Tables

Nomenclature

Variables

a	Nominal decline rate, dimensionless
b	Arps decline exponent, dimensionless
d_q/d_t	Slope, dimensionless
D	Decline constant, d^{-1}
D_i	Initial decline constant, d^{-1}
D_{lim}	Decline rate, d^{-1}
D_∞	Decline rate constant over a long period, d^{-1}
EUR	Estimated Ultimate Recovery, bcf
f	Non-linear function, dimensionless
G_p	Cumulative gas production, bcf
K	Carrying constant, dimensionless
L_t	Linear component
m	Slope, dimensionless
$n/\hat{n}$	Time exponent, dimensionless
n	Exponent parameter of SEPD model, dimensionless
N_p	Cumulative gas production, bcf
N_t	Non-linear component, dimensionless
NPV	Net present value, currency
OPEX	Operating expenditure, currency

p	Number of input nodes, dimensionless
q	Number of hidden nodes
$q/\widehat{q_\iota}$	Production rate, Mscf/day or STB/day or m^3/day
q_i/q_o	Initial production rate, Mscf/day or STB/day or m^3/day
q_∞	Production rate at infinite time, m^3/s
t	Time, days
$t(a.m)$	Time constant, 1/s
Y_t	Time series data, days
r_i	Advances of the pressure drawdown front, ft
c_t	Compressibility, $m^{2/n}$
f_t	Forget gate
i_t	Input gate
O_t	Output gate
c_t	Long-term state
$\tilde{C}_t$	New candidate value

Greek Symbols

ε	Random shock
$\propto_1$	Self-regression coefficient
$\propto_j/\beta_{ij}$	Parameters of the model
β_i	Late life period constant, dimensionless
β_e	Early period constant, dimensionless

τ	Characteristic time parameter for SEPD model, day
$\phi(z)$	Porosity, g/cm^2
$\kappa(z)$	Permeability, H/m
μ	Viscosity, Pa.s

Chapter 1

Background and Motivations

1.1. Introduction

In recent years, shale gas reservoirs have gradually become the major sources of natural gas production around the world (Zhang et al., 2017). Shales and silts are the most abundant sedimentary rocks in the earth's crust and from recent years' activities shale gas will constitute the largest component in gas production globally, as conventional reservoirs continually decline (Wang, 2017). According to Wang (2017), unlike conventional reservoirs, the shale gas reservoirs tend to be more expensive to develop and require special technologies to enable the gas to be produced at an economical rate, due to its extremely low matrix permeability and porosity. Thus, modelling the shale gas production and its declines is essential to predict how fast the gas can be produced and be turned into revenue from each well, and the economic viability of producing natural gas from the operated shale plays (Wang, 2017).

Rate-time decline curve extrapolation is one of the oldest and most commonly used tools used by a petroleum engineer (Fetkovich, 1980). Results obtained for a well are subject to a wide range of alternate interpretations, mostly as a function of the experience and objectives of the evaluator. Recent efforts in the area of decline curve analysis (DCA) have been directed toward a purely computerized statistical approach, its basic objective being to arrive at a unique "unbiased" interpretation (Fetkovich, 1980). In the past few decades, several DCA models have been proposed and benchmarked with commercial reservoir simulators or shale gas production data before being applied to more shale gas reservoirs (Tan et al., 2018).

For these reasons, this research is intended to assist the petroleum industry worldwide and particularly in South Africa in understanding the decline behavior and characteristics of shale gas reservoirs by:

- Analyzing field production data using a developed decline curve model.
- Validating the model against existing models, and
- Determining which model should receive more focus and attention.

Secondly, this research is to determine the financial impact of decline curve analysis on the present and future value of shale gas reservoirs.

1.2. Problem Statement

Numerous studies have highlighted the importance of DCA models. However, there are limitations. Analysis conducted using these techniques for the prediction and estimation of reservoirs in shale well production with the decline exponent, b > 1 have highlighted shortcomings in the models (Paryani et al., 2018). There can be an under-estimation, or an overestimation of the estimated ultimate recovery (EUR) of reserves. Considering these facts, there is scope for developing an improved model, which addresses these shortcomings. The shortcomings are listed below for the different DCA models.

- **Arps Hyperbolic Decline:** post-production overestimation.
- **Modified Hyperbolic Curve:** is still unable to determine D_{lim} for production data.
- **Power Law Exponential Decline:** there are four unknown variables to solve.
- **Stretched Exponential Decline:** requires sufficiently long production times.
- **The Extended Exponential Model:** the parameter β_l has an incomplete influence on the curve fitting and is therefore fixed.
- **Duong's Decline:** for extended periods, a proper rate initialization against pressure is required, and in the event of water breakthrough, *a* and *m* increases.
- **Logistic Growth:** growth is only possible up to a certain size.

1.2.1. Main Focus

The purpose of this study is to develop a decline curve model, which can accurately forecast or predict shale gas production when exploration begins in South Africa. The reason for conducting this investigation is to determine whether improved hybrid models can assist in guiding petroleum companies to recover from loss of revenue.

1.2.2. Sub-Focus

The research is broken down into two separate sub-foci:

- The first sub-problem is to develop a decline curve model which can accurately predict or forecast future production from field production data.
- The second sub-problem is to determine how decline curve analysis can assist in future revenue estimation.

1.3. Research Aim (s) and Objective (s)

The main aim of this study is to model the decline of shale gas reservoirs mathematically, and determine its financial impact. The developed model will be validated against three DCA models, in order to assess the level of accuracy in the prediction results.

To meet the aims of this research five main objectives will be investigated:

- Analyse the accuracy and sensitivity of Decline Curve Models.
- Develop hybrid decline curve models.
- Validate the developed model against other DCA models i.e. the three models that have been identified for this study.
- Determine which of the decline models accurately forecasts the production decline of shale gas reservoirs (estimate the quantity of gas available), and
- Lastly, determine the present and future value of the reservoir.

1.4. Delimitation (s) and Limitation (s) of the Study

1.4.1. Delimitation (s)

The research was subject to the following delimitation; while the results and recommendations of this research study could be applicable to most shale wells, the research data used was primarily from United States shale wells.

1.4.2. Limitation (s)

- Production data from previous studies was extrapolated and used in the study. Data inputs such as volume, temperature and pressure were not available.
- Due to the limitation on input data, a univariant ANN model was developed and tested.

- Lastly, for the purposes of this study, due to the limited information available regarding the equity investment, depletion allowance, total net operating expenses and severance and ad valorem expenses, these parameters will be excluded from the study and the financial model will only consider the estimated total net revenue.

1.5. Thesis Organisation

The thesis is organised into nine chapters. The first chapter covers the introduction to this study and highlights the background to the study, the problem statement, the objectives and the organisation. Chapter 2 deals with the review of literature and focuses on previous research. Chapter 3 discusses the research methodology. Chapter 4 discusses the evaluation and sensitivity analysis of decline curve models. Chapters 5, 6 and 7 examines model development, simulations and validations, respectively. Chapter 8 of the thesis briefly discusses the financial modelling analysis. Lastly, Chapter 9 covers the conclusions and recommendations.

Chapter 2

The Literature Review

2.1. Introduction

In recent years, shale gases from reservoirs have gradually become the major sources of natural gas production around the world (Zhang et al., 2017). Shale gas as an energy resource is growing rapidly and currently constitutes more than 20% of the drilled gas production in the United States (US) (Knudsen et al., 2012). Exploitation of shale gas is land-based, and normally requires a very large number of wells to achieve profitable recovery rates (Knudsen et al., 2012). Unlike conventional reservoirs, the shale gas reservoirs tend to be more expensive to develop and require special technologies to enable the gas to be produced at an economical rate, due to its extremely low matrix permeability and porosity (Wang et al., 2017).

Shale gas reservoirs have been analysed around the world, however, development has been stalled due to a variety of issues one of which is an accurate production forecast (Nwaobi and Anandarajah, 2018). Estimated ultimate recovery (EUR) approximates the quantity of oil or gas that is potentially recoverable or has already been recovered from a reserve or well. EUR is similar in concept to recoverable reserves (Nwaobi and Anandarajah, 2018). The EUR is an essential factor for both investors and policy makers in appraising petroleum resources (Nwaobi and Anandarajah, 2018).

Thus, modelling the shale gas production and its declines correctly is essential to predict how fast the gas can be produced and turned into revenue from each well, and the economic viability of producing natural gas from the operated shale plays (Wang et al., 2017).

2.2. Unconventional Reservoirs: Overview of Shale Gas Production

According to Moridis et al. (2013), unconventional gas resources (UGRs) is a term used to describe accumulations that are currently hard to characterize and commercially produce using traditional exploration and production technologies (Moridis et al., 2013). They include:

- Natural gas from shale formations (shale gas).
- Natural gas from coal seams (coal-bed methane).
- Crude oil from shale formations or other formations with low permeability (tight or shale oil).

Due to the growing demand for energy and the declining of conventional oil and gas production, shale gas has received increasing attention worldwide. Compared with other fossil fuels, shale gas is a clean-burning and efficient energy resource. **Figure 2.1** provides a list of the 10 countries holding the largest resources of shale gas (Yuan et al., 2015).

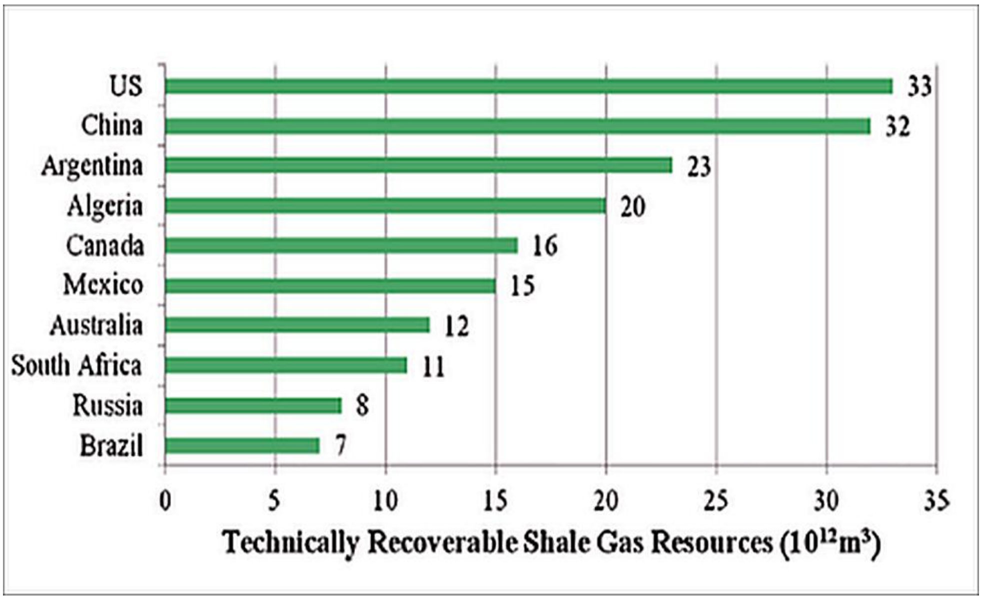


Figure 2.1: Top 10 countries holding the largest shale gas resources (Yuan et al., 2015)

The influence of shale gas usually needs a sizeable quantity of wells to achieve favourable recovery rates (Knudsen et al., 2012). Nwaobi and Anandarajah (2018) explained that shale gas reservoir production and viability have been investigated globally but progress has been slow due to a number of concerns, one of which is a precise production forecast. Nwaobi and Anandarajah (2018) went on to define that the quantity of shale gas reserve that can be recovered is the EUR for the petroleum industry. The EUR is a key factor for stakeholders and policymakers in evaluating petroleum resources (Nwaobi and Anandarajah, 2018).

Shale gas has played a key role in giving the US, formerly a natural gas importer, the ability to export natural gas. This development has helped the US to successfully ensure its energy security and significantly reduce carbon emissions. Canada has become the second country to achieve the commercial exploitation of shale gas. The successful development of shale gas in Canada, has injected new vitality into the nation's natural gas production, which had previously experienced a rapid decline (Yuan et al., 2015). A new report from the International Energy Agency says that a remarkable ability to unlock new resources cost-effectively will push US oil and gas output to a level of 50 percent higher than any country has ever managed before. That is expected to have a considerable impact on both North America and the world, with the expansion set to reorder international trade flows as well as fuelling investment in energy-intensive industries such as petrochemicals (McCarthy, 2017).

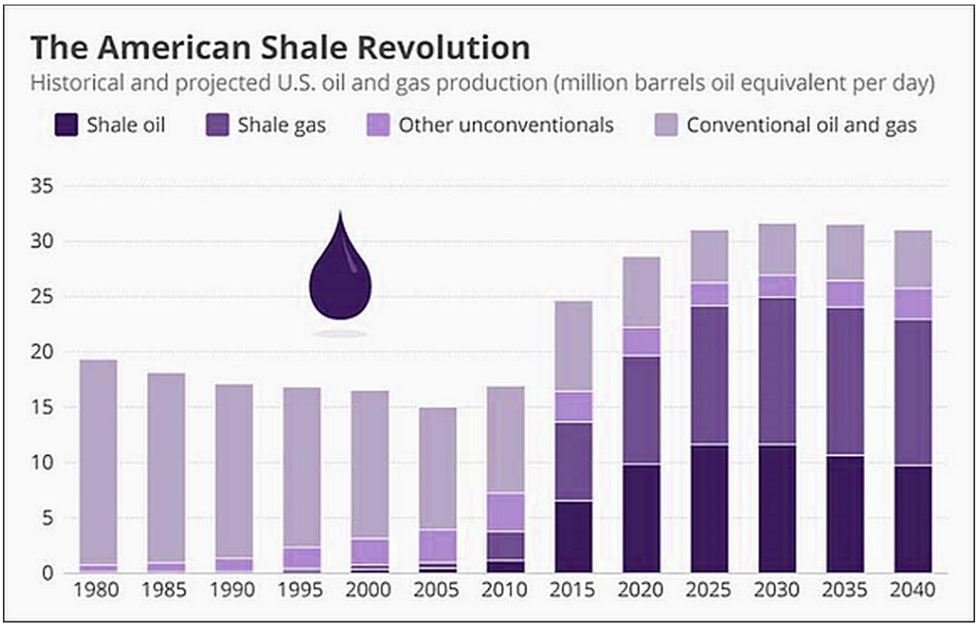


Figure 2.2: Shale gas production growth in the US from 1980 and projected to 2010 (McCarthy, 2017)

2.3. Characteristics and Production Behaviour of Shale Gas

Shale gas reservoirs possesses characteristics such as ultra-low permeability, a no trap mechanism, and the gas is tightly absorbed to the rock particle, which is the opposite of a conventional reservoir (Adekoya, 2009). Hydraulic fracturing is often used in reservoirs with low permeability that are not able to reach economic production rates (Adekoya, 2009). This is very different in character to the naturally fractured reservoirs that are classified as having dual porosity (Adekoya, 2009). There are four different flow regimes that can occur in a hydraulically fractured reservoir and several flow periods can exist during the life cycle of a shale gas well (Adekoya, 2009; Nelson, 2009). These comprise fracture linear flow, fracture boundary flow, matrix linear flow, and lastly matrix boundary flow (Joshi, 2012). Joshi (2012) explained the different flow regimes for shale gas reservoirs as follows:

- **Fracture**/**Early Linear Flow**: Is a transient flow regime (TFR) that occurs when the production flow is linear to the single fractures. This flow regime governs the known life of most shale wells. A negative half slope on a log-log plot of rate versus time can be used to differentiate this linear flow.

- **Fracture Boundary Flow**: Follows after a certain period of production when an interference occurs i.e., from linear to simulated reservoir volume (SRV). Many of the existing horizontal shale wells have not experienced this regime, but some of the newer wells with huge fracture treatments have been observing this regime early. This can be observed on a log-log plot by deviation from a –1/2 slope line on a log-log plot of rate versus time.

- **Matrix Linear Flow**: When production from the matrix, beyond the SRV, starts to govern the production, a linear type of flow will be seen. This regime will probably not be observed in the economic life of the well. Comparable to fracture linear flow, this regime can be observed using a negative half slope line on a log-log plot of rate versus time.

- **Matrix Boundary Flow**: After the outer matrix transient has reached the drainage boundaries of the well, a deviation from the negative half slope, corresponding to matrix linear flow, will be observed. This deviation is equivalent to matrix boundary flow. Similarly, to the matrix, a linear flow will probably not be observed.

Figure 2.3 depicts the different flow regimes which has been explained in a typical multistage fractured horizontal well in the Marcellus Shale Well (Joshi, 2012). The figure was created using the Fekete Well Test Software and using input properties obtained from literature (Joshi, 2012).

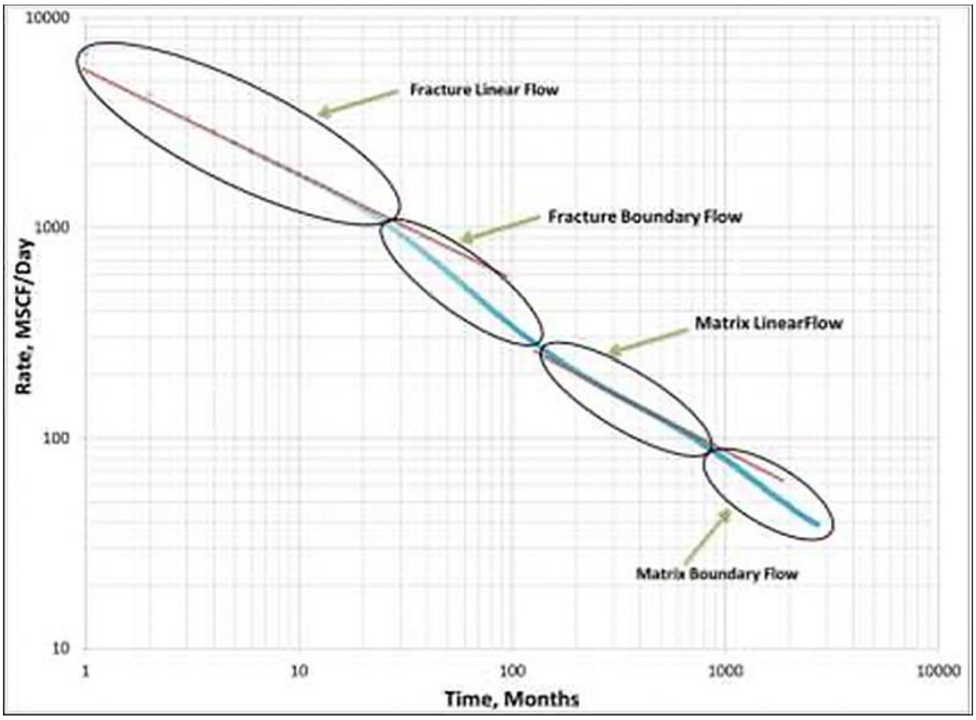


Figure 2.3: Flow regimes in a multistage fractured horizontal well (Joshi, 2012)

2.4. Characteristics and Production Behaviour of Shale Gas

2.4.1. Shale Gas in the United States

The US has abundant natural gas resources within the Barnett Shale Well, Haynesville/Bossier Shale, Antrim Shale, Fayetteville Shale, New Albany Shale, and Marcellus Shale Well (Kargbo et al., 2010). At the annual production rate of about 19.3 Tcf, there is enough natural gas to supply the US for the next 90 years, with some estimates extending the supply to 116 years. The total number of natural gas and condensate wells in the U.S. increased by 5.7% in 2008 to a record of 478,562 with some of the produced natural gas lost via flaring. The Barnett Shale Well is located in the Forth-Worth basin of Texas and was the first modern commercial shale play in the US, having been discovered in 1981 (Kenomore et al., 2018). Until 2012, it was the largest shale gas basin (Newark East Field) in the world, before been replaced by Marcellus. Natural gas extraction in the Marcellus Shale Well is currently an expensive endeavour. A typical horizontal drilled well, using multistage fracturing techniques, costs roughly $3-5 million to complete. The large amount of water used, and management of the wastewater are also very costly factors. According to the Global Shale Gas Initiative (GSGI), there are more than 688 shale deposits worldwide in 142 basins and 48 major shale basins which are located in 32 countries (Wang, 2017). **Figure 2.4** shows the shale basin across the United States (Kargbo et al., 2010).

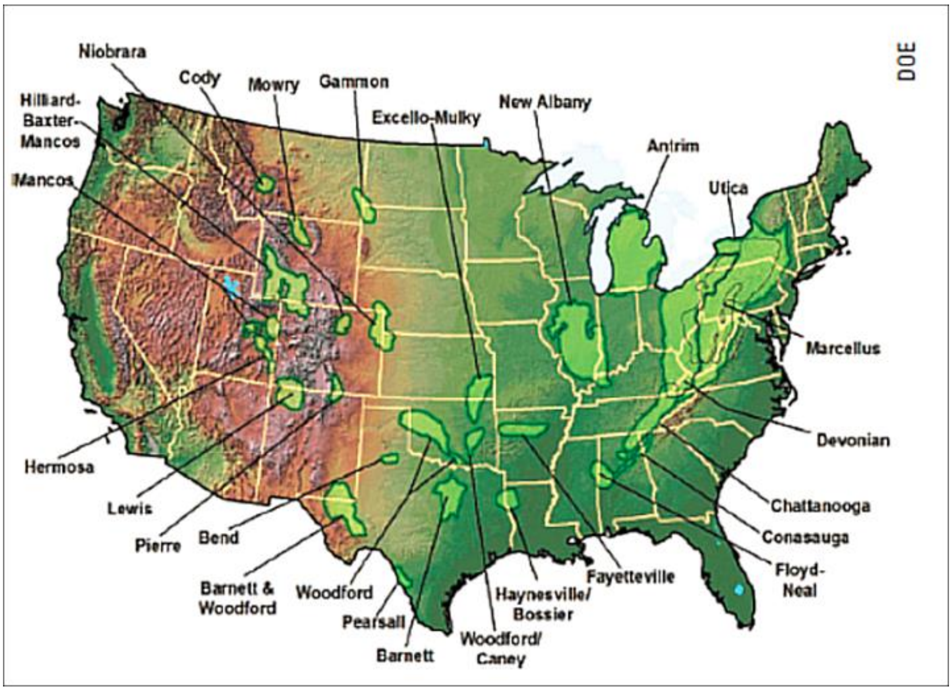


Figure 2.4: Shale basins across the US (Kargbo et al., 2010)

2.4.2. Hydraulic Fracturing

To simulate gas production in shale, complex gas transport mechanisms and the presence of natural and hydraulic fractures need to be considered, which makes the simulation of shale gas reservoirs a challenging task (Tan et al., 2018). Hydraulic fracturing was developed in the US in the late 1940s to assist in the stimulation of oil and natural gas wells (Steyl et al., 2012). The technique itself is mechanically related to three other phenomena, which have previously been well documented. These are:

- pressure parting in water injection wells in secondary-recovery operations,
- lost circulation during drilling, and
- the breakdown of formations during squeeze-cementing operations, all of which appear to involve the formation of open fractures by pressure applied in a wellbore (Hubbert and Wallis, 1972).

The fracking process (**Figure 2.5**) is very technical and intensive. During the drilling process, each well can require up to 6 million gallons of water to reach the desired output. Drilling companies typically pipe the water they use from nearby rivers and streams.

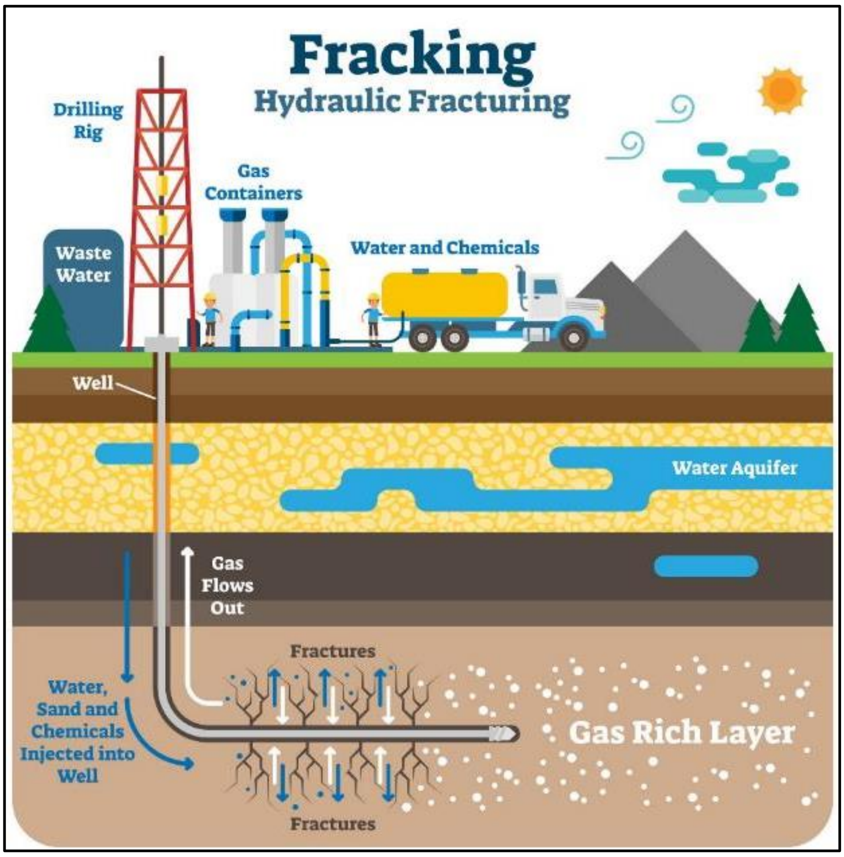


Figure 2.5: The hydraulic fracturing process (Kendon, 2019)

The following explains the different processes - as seen in the figure above:

1. A mixture of water, sand and chemicals is shot down the well at high pressure.
2. The pressurized mix causes fissures to develop. The sand in the mixture helps keep the fissures open, allowing oil to seep into the well.
3. The seeping oil is then pumped back up the well.

2.5. Shale Gas Reservoirs in South Africa

The announcement that gas exploration will take place in the Karoo has raised both concern and elation. The excitement is partly due to the creation of employment prospects while the concern is driven by a fear that the pristine way of life in the Karoo area might be destroyed (Steyl et al., 2012). One major concern that has been raised is the influence of hydraulic fracturing on groundwater resources and the top geological strata in the Karoo. Due to the present energy shortfall in South Africa, the requirement for new energy sources has gained new momentum and part of this new focus is on shale gas in Karoo type formations. A number of companies are planning to explore and exploit the shale gas reserves of the Karoo of South Africa and elsewhere, with shale gas projects, have encountered substantial local opposition (Wait and Rossouw, 2014). The most interesting aspect of this is that the area available for natural gas development is substantially larger than just the Karoo, with exploration areas covering six of the nine provinces in South Africa. A five-spot pumping test in the Waterberg has been operated since 2004 by Anglo Operations and 20 boreholes have been drilled in the main Karoo since the beginning of 2008 to test for coal-bed methane production potential (Steyl et al., 2012).

2.6. Shale Gas Extraction: Environmental Impact

Shale gas exploration and exploitation require proper guidelines in view of the environment (Dayal, 2017). Shale gas exploration includes the shallow seismic study of the basin, either explosives or Vibroseis are used for obtaining the seismic data (Dayal, 2017). According to Soomro et al. (2017), the following aspects affect the environment, due to shale gas extraction:

1. Surface and Ground Water Contamination

In the high risk of water contamination at different steps of the well site preparation, fracking process, fluids flow toward the surface, and contamination during well abandonment on the surface at ground level has been observed. Soomro et al’s (2017) study considered the water contamination risks as being caused by the following:

- Poor well design or poor casing structure.
- Well, kick or mechanical equipment failure.
- Movement of combustible natural gas towards the water storage/supplies.
- Geological conditions are inadequate.
- Inadequate planning and site preparation and management.

2. Air emissions

According to Theodori (2013), the natural gas produced during the fracking operations can be bad for the atmosphere which can be compared with the effects of coal usage. This is due to the discharge into the atmosphere. Loh and Loh (2016) suggested that the emissions from shale gas are between 20-100% more than that of coal during the life cycle of greenhouse gas on a 20-year timeframe basis. The emissions produced from this fracking operation are harmful for the environment.

3. Land and Take

Certain aspects of Soomro et al's (2017) study showed that the land used in shale gas extraction has a significant risk of impact because of the small size of land required for production stage of coal, compared to the large size of the land required for the fracking process.

4. Noise Pollution

There are many sources of noise in shale gas extraction i.e., during excavation, installation and drilling, running generators, processing and transport and the level varies from stage to stage i.e., during the preparation and production cycle (Soomro et al., 2017).

5. Seismicity

The seismic effect associated with the fracking process is negligible up to a magnitude of 3 on the earthquake magnitude scale which is undetectable. Therefore, the risk attached to this hazard is low (Howarth et al., 2011). The summary of risk level at various stages of field development to individual site can be seen in **Table 2.1**.

Table 2.1: Level of risk at various stages of field development (Soomro et al., 2017)

Aspects of environmental hazards		Site start-ups	Well Design	Fracturing Operations	Well Completion	Production	Well Abandonment (Pre and Post)	Overall Rating
Water Contamination	Ground	N/A[2]	Low	Moderate - High	High	Moderate - High	N/C[3]	High
	Surface	Low	Moderate	Moderate - High	High	Low	N/A	High
Resources of water		N/A	N/A	Moderate	N/A	Moderate	N/A	Moderate
Air Impact		Low	Moderate	Moderate	Moderate	Moderate	Low	Moderate
Land and take impact		Moderate	N/A	N/A	N/A	Moderate	N/C	Moderate
Seismicity Impact		N/A	N/A	Low	Low	N/A	N/A	Low

[2] Not Applicable (N/A) - Impact not relevant to this stage of development
[3] Not classifiable (N/C): Insufficient information available for the significance of this impact to be assessed

2.7. Modelling Shale Gas Reservoirs

Reliably predicting the long-term production performance of shale (unconventional) reservoirs has been challenging. The petroleum industry needs simple, easily applied and rapid methods of forecasting production and estimating reserves. Therefore, an empirical technique such as DCA is an appealing alternative compared to reservoir simulation and applied analytical methods (Makinde and Lee, 2017). Due to the relative simplicity of DCA, it is considered the most common method used in the industry.

DCA is a technique where production data from a well or reservoir is used to predict the well/reservoir's future production (Paryani et al., 2018). Two important goals of DCA are to estimate the remaining reserves and their remaining life, down to a specified economic limit, both of which are important for determining the economic viability of a shale resource play (Chen et al., 1992). In addition, no rate decline model can be expected to provide a unique forecast of future performance or EUR (Leblanc and Okouma, 2018).

Production data from unconventional reservoirs exhibits extensive periods of transient (linear and bilinear) flow behaviour, due to low/ultra-low permeability. This often leads to an over-estimate of gas-in-place/reserves when using the conventional rate-time relations (exponential and hyperbolic rate-time relations). A common practice for reserves estimation in unconventional reservoirs is to use the hyperbolic relation and constrain the ultimate extrapolation by including a terminal exponential decline value, or a "modified hyperbolic" designation. In an attempt to better represent the general character of production data for multi-stage, fractured horizontal well in an ultra-low permeability reservoir, numerous authors have developed rate-time models using specific assumptions to best represent a particular scenario (Leblanc and Okouma, 2018).

2.7.1. Arps Decline Curve and the Modified Hyperbolic Decline Model (MHD)

Arps decline curve analysis is the most commonly used method of estimating ultimate recoverable reserves and future performance (Boah et al., 2018). Paryani et al. (2018) reasons this to be a reliable history match (even with $b > 1$) and its simplicity. The modelling process is based on vital assumptions: that past operating conditions will remain unaffected, a well is produced at or near capacity, and the well's drainage remains constant and is produced at a constant bottom-hole pressure (Ali and Sheng, 2015).

Notably, the Arps model is only applicable in pseudo-steady flows when the flow regime transfers from linear flows to boundary-dominated flows (BDF) (Yuhu et al., 2016). This indicates that the Arps Equations are not applicable to the production forecasting of the entire decline process of horizontal wells in low-permeability reservoirs (Li et al., 2018). The Arps DCA can be divided into three types: exponential Equation (2.1), hyperbolic Equation (2.2), and harmonic Equation (2.3) (Arps, 1945; Qu and Lin, 2018) (**Figure 2.6**).

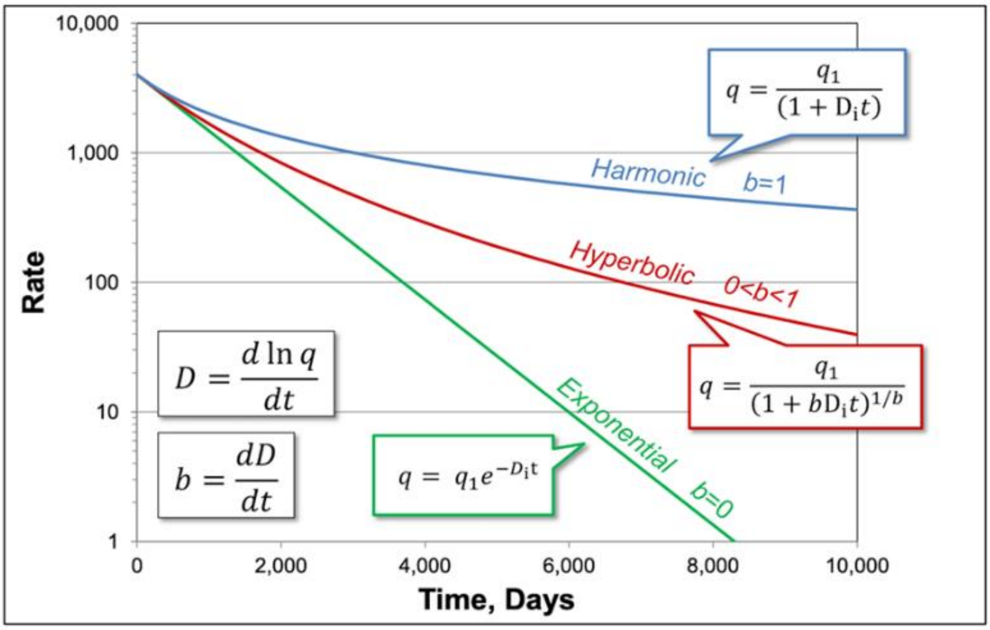


Figure 2.6: Arps three decline models and formulas on a semi-log plot (Kanfar, 2013)

$$q = q_i e^{-Dt} \tag{2.1}$$

$$q = \frac{q_i}{(1 + bD_i t)^{\frac{1}{b}}} \tag{2.2}$$

$$q = \frac{q_i}{1 + D_i} \tag{2.3}$$

where q is the flow rate in STB/day or Mscf/day, q_i is the initial flow rate in STB/day or Mscf/day, D is the decline constant while D_i is the initial decline constant, both of which are measured in days $^{-1}$, and b is the decline exponent.

The most commonly employed hyperbolic form of Arps decline Equation (2.2) is used for shale reservoirs (**Figure 2.7**).

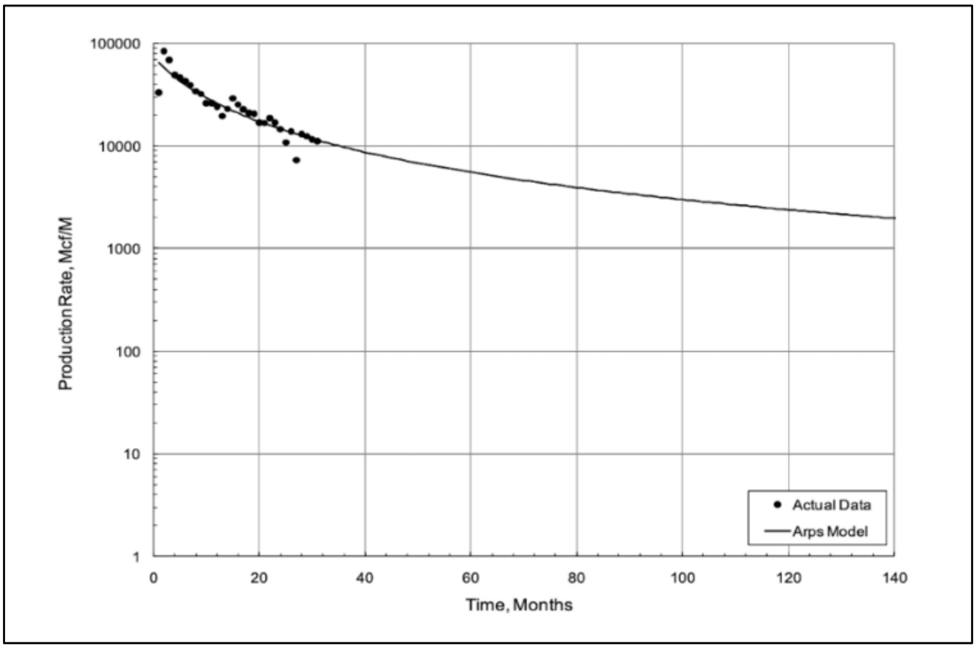


Figure 2.7: Production data for a gas well fit with the Arps hyperbolic model (Clark, 2011a)

The hyperbolic decline Equation is suitable to use due to the "best fit" that it provides for the long transient linear-flow regime observed in shale gas wells with b values greater than unity (Qu and Lin, 2018). The model results in post-production overestimation due to the decrease in the decline rate with production time. Due to the overestimation. Robertson (1988) suggested a revised version of the hyperbolic decline model for shale gas production decline. The equation is given as:

$$q = \frac{q_i}{(1 + nD_i t)^{1/n}} \quad (D > D_{lim}) \qquad \textbf{(2.4)}$$

$$q = q_i \; exp(-D_{lim} t) \quad (D \leq D_{lim}) \qquad \textbf{(2.5)}$$

where q is the production rate in m^3/d or STB/day, D_{lim} is the decline rate in d^{-1}, and n is the time exponent. They suggested that the hyperbolic decline model sometimes yields unrealistically high reserve estimates. They made an assumption that the rate of decline starts at 30% of flow and usually declines in a hyperbolic way (Robertson, 1988). This modified model considers when the hyperbolic decline in the early life of a well transfers to exponential decline in its later life (Robertson, 1988). The switching process can be determined by applying computer programs. The switching point is when the decline rate is smaller than a certain limit (usually 5%) (Robertson, 1988). The Modified Hyperbolic Decline (MHD) model addresses the overestimation limitation of EUR. However, it is still unable to determine D_{lim} for production data (Yuhu et al., 2016).

2.7.2. Power Law Exponential Model (PLE)

Ilk et al. (2008) presented the PLE, which is an extension of the exponential Arps formula for the decline degree in shale reservoirs. This model was developed precisely for shale gas reservoirs (SGR) and approximates the rate of decline with a power law decline. The PLE model matches production data in both the transient and boundary-dominated regions, without being hypersensitive to remaining reserve estimates (McNeil et al., 2009). Seshadri and Mattar (2010) presented that the PLE model can model transient radial and linear flows, while Kanfar and Wattenbarger (2012) proved that the model is reliable for linear flow, bilinear flow and linear flow, followed by BDF, or bilinear flow followed by linear flow and finished with BDF flow. Vanorsdale (2013) deduced that when the flow regime changes throughout the initial 10 years of the well, the PLE model would yield a very optimistic recovery. The model characterizes the decline rate by infinite time, D_∞ which is defined as a “loss ratio” (which is assumed to be constant from Arp) (Li et al., 2018). The production rate is derived as follows:

$$\frac{q}{dq/dt} = -b \quad \textbf{(2.6)}$$

$$b = D_{\infty} {}_{+} D_i t^{-(1-\hat{n})} \quad \textbf{(2.7)}$$

where dq/dt is the slope, D_∞ is the decline rate over a long-term period, and $\hat{n}$ is the time exponent. By substituting the above equations, the production rate is obtained:

$$q(t) = \hat{q}_i e^{[-D_\infty t - \hat{D}_i t^{\hat{n}}]} \qquad \textbf{(2.8)}$$

In this model, there are four unknown variables: $\hat{q}_i, \hat{D}_{i,}\ D_\infty\ and\ \hat{n}$, which result in several degrees of freedom and may be clumsy to use or solve (Hu et al., 2018). According to Johnson et al. (2009), the $D\infty$ parameter is difficult to determine. However, there are advantages to this model in that the extra variables allow for both transient and boundary flow, and the equation for production rate seems comparable to the Arps exponential equation (Paryani et al., 2018).

2.7.3. Stretched Exponential Decline Model (SEDM)

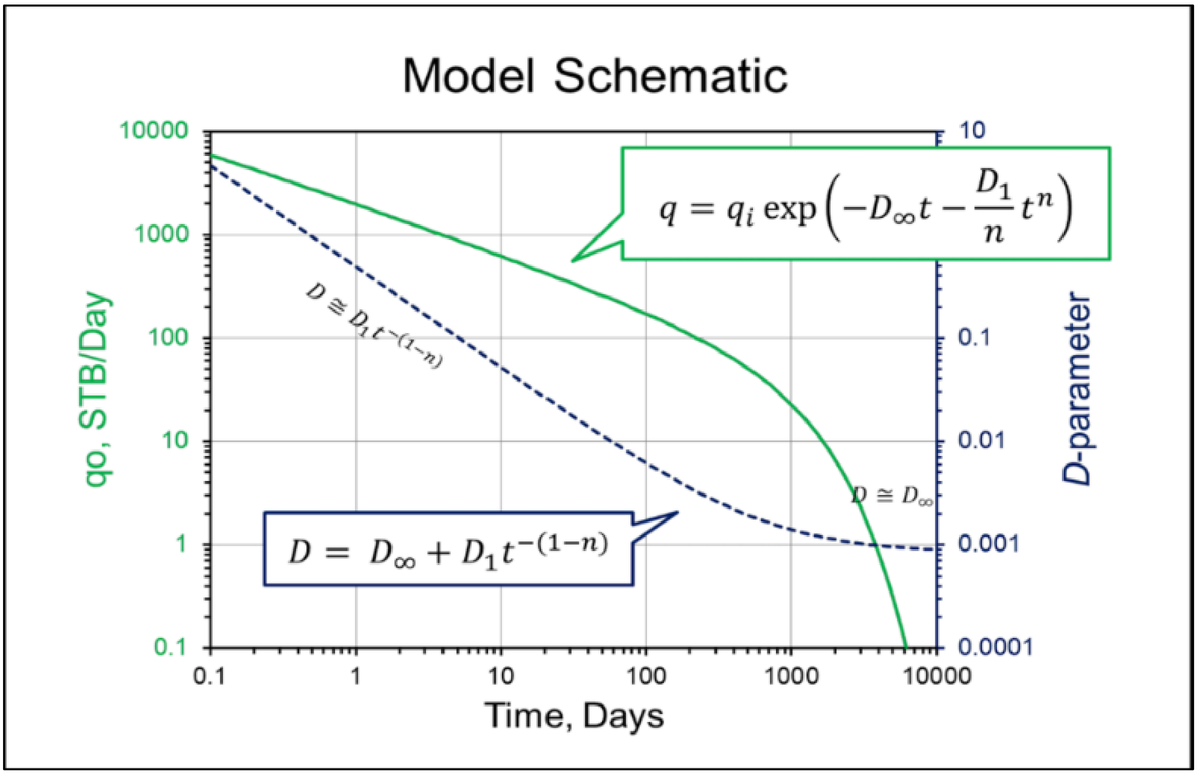


Figure 2.8: PLE schematic developed by Ilk et al. (2008) (Kanfar, 2013)

Valkó (2009), Valkó, and Lee (2010) applied the SEDM in shale wells, which is an empirical method that differs from the Arps Equations, as it describes the decline trend of production data obtained from unconventional reservoirs. It was developed to fit transient flow regimes (Joshi, 2012; Kanfar and Wattenbarger, 2012).

The significant advantages of the model are the bounded nature of EUR without limits on time or rate, and the straight-line behavior of a recovery potential expression

(Valkó, 2009). The model differs from other models, since it does have a basis in physics and is directed by a major differential equation (Ali and Sheng, 2015). It is used to model aftershock decay rates (Valkó, and Lee, 2010). The production rate declines with time, according to the following equations:

$$\frac{dq}{dt} = -n\left(\frac{t^n}{\tau}\right)\frac{q}{t} \qquad \textbf{(2.9)}$$

$$q = q_i \exp\left[-\left(\frac{t}{\tau}\right)^n\right] \qquad \textbf{(2.10)}$$

$$Q = \frac{{q_i}^n}{n}\left\{\Gamma\left[\frac{1}{n}\right] - \Gamma\left[\frac{1}{n}.\left(\frac{t}{\tau}\right)^n\right]\right\} \qquad \textbf{(2.11)}$$

$$\mathrm{EUR} = \frac{{q_i}^\tau}{n}\,\tau\left[\frac{1}{n}\right] \qquad \textbf{(2.12)}$$

This method defines a characteristic number of periods, τ and a dimensionless exponent, *n*, of the ratio of time, *t*. It also uses observed cumulative production along with theoretical cumulative production, derived from the integral of the rate-time equation to estimate remaining technically recoverable volumes. Equation (2.10) appears similar to the PLE model; however, it differs, as it does not rely on a single interpretation of parameters. Instead, it uses two-parameter gamma functions (Johnson et al., 2009).

In addition, there are no single τ and *n* parameters, but instead, a sum of multiple exponential declines, which follows the fat tail distribution (Valkó, 2009). SEDM requires an iterative process to determine the value of the parameter, *n*. The model can only estimate the recoverable volumes with an abandonment rate of zero, as opposed to commercial volumes with economic cut-off rates and it has not been widely used (Can and Kabir, 2011). However, Can and Kabir (2011) showed that in tight formations where transient flow period is extremely long, the SEDM has been successful in modeling the rate-time behavior and provides more realistic reserve estimates compared to Arps decline relations.

2.7.4. The Extended Exponential Model (EEDM)

Zhang et al. (2016a and 2016b) presented a renewed experimental method, the EEDM, as a simple formula to forecast shale oil and gas well performance. They proposed a mechanism of “growing drainage volume” to conceptualize and model the performance of shale wells. This model combines the exponential decline equation proposed by Fetkovich (1980) Equation (2.13) with the derived empirical Equation (2.14). The EEDM includes both transient and BDF flow in a single equation, and it can match the historical data with a smooth curve throughout the transition period from transient to BDF flow regimes. Furthermore, the model is simple and can easily be applied (Zhang et al., 2016). It is also able to project the future production by fitting all of the historical production data from the beginning of the production decline.

Paryani et al. (2018) stated that the model contains two decline constants and a decline exponent. Particularly noteworthy is the fact that the production data fits using a smooth curve through the whole flow systems (Li et al., 2018). The advantage of the model is that both early and late production profiles can be captured once $\beta_e\ and\ \beta_l$ have been calibrated, using the production data (Zhang et al., 2016a and 2016b). However, as parameter β_l has an incomplete influence on the curve fitting, it is therefore fixed:

$$q = q_i e^{-at} \quad \textbf{(2.13)}$$

$$a = \beta_i + \beta_e \quad \textbf{(2.14)}$$

where *a* is the nominal decline rate, β_i is the late-life period constant, and β_e is the early period constant. Combining Equations (2.13) and (2.14) and taking the logarithm of each side, the equation below (the exponential decline equation) is obtained:

$$\frac{In\frac{q}{q_o}}{t} = \beta_l + \beta_e e^{-t^n} \quad \textbf{(2.15)}$$

where q_o is the initial production rate in m^3/s. **Figure 2.9** shows decline curves which are controlled by Zhang’s constants β_e and $\beta_{i.}$.

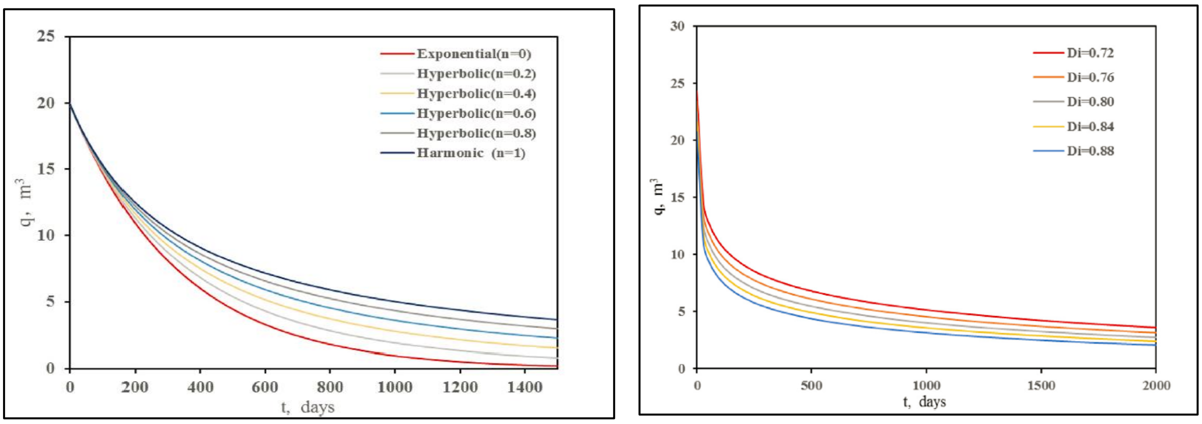


Figure 2.9: Zhang's decline curves controlled by β_e and β_i

2.7.5. Duong's Decline Model

Duong (2011) presented an unconventional rate decline method to evaluate the performance of shale gas wells that does not depend on the fracture types. **Figure 2.10** summarizes the computation of the Duong's Model.

The model assumes linear or near-linear flow, as indicated by a log–log plot of rate over cumulative production versus time, which yielded a straight-line tendency (Lee and Kim, 2016). The rate is calculated in the model using the following Equation (2.16):

$$(t) = q_i t(a,m) + q_\infty \qquad \textbf{(2.16)}$$

where $t\,(a,m)$ is the time constant in 1/s, and q_∞ is the production rate at infinite time in m^3/s. The cumulative production and time constant are calculated as:

$$G_p = \frac{qt(a,m)}{at^{-m}} \qquad \textbf{(2.17)}$$

$$t(a,m) = t^{-m}\exp(\frac{a}{1-m}(t^{1-m}-1)) \qquad \textbf{(2.18)}$$

where G_p is the cumulative gas production in bcf and m is the slope.

Paryani et al. (2018) indicated the key restrictions of the model are:

- Firstly, if a well is closed for extended periods, a proper rate initialization against pressure is required to obtain precise values of parameters *a* and *m* and,

- Secondly, that in the event of water breakthrough, there is a sudden decrease in the decline rate, which then causes an increase in the values of the *a* and *m* parameters.

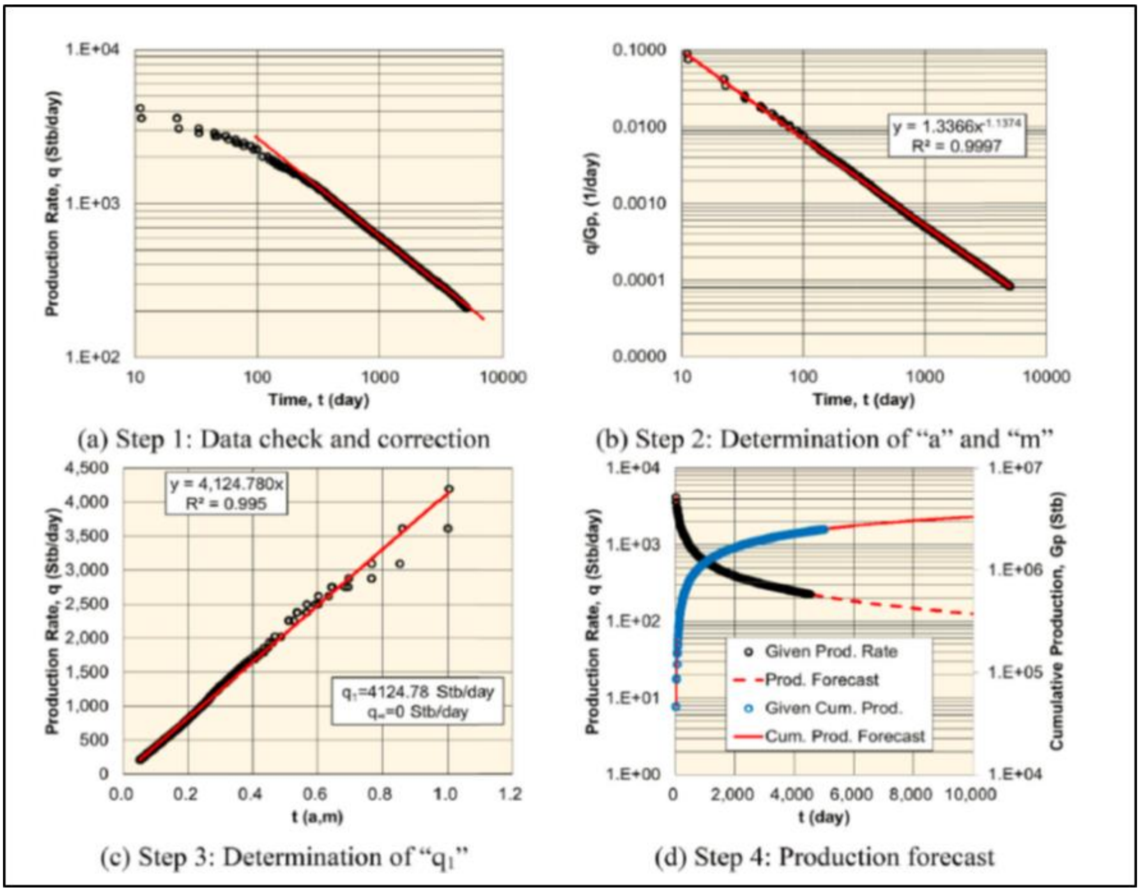


Figure 2.60: Four steps of computing the Duong's model (Hu et al., 2018)

Vanorsdale (2013), found like in the case of the PLE model, the Duong's model also yielded a very optimistic recovery when the flow regime changes throughout the initial 10 years. The author went on to indicate that the model might provide conservative recovery estimates in vertical, non-hydraulic fractured classical shale wells (Vanorsdale, 2013). However, Lee et al. (2016) indicated that the Duong's model appeared to fit field data from various shale plays quite well and provided an effective alternative to Arps hyperbolic model.

2.7.6. Logistic Growth Model (LGM)

Logistic Growth Models developed belong to a group of mathematical models used to forecast growth in numerous applications (Lee and Kim, 2016) and were previously used to model population growth (Clark, 2011a; Clark et al., 2011b). They were developed to forecast reservoirs with extremely low permeability (Hu et al., 2018). LGM is very flexible and confident in modelling long transient boundary-dominated performances of unconventional reservoirs (Li et al., 2018). The model incorporates known physical volumetric quantities of oil and gas into the forecast, to constrain the reserve estimate to a reasonable quantity. LGM is capable of trending existing production data and providing reasonable forecasts of future production. The logistic growth model does not extrapolate to non-physical values (Clark, 2011a).

Tsoularis and Wallace (2002) discussed a development in this regard by Bacaër and Verhulst (1838), who considered that for the population model, a steady population would consequently possess a saturation level characteristic, typically called the carrying capacity, *K*, which forms a numerical upper bound on the growth size. In order to include this limiting characteristic, they introduced the logistic growth equation as an extension to the exponential model (Clark et al., 2011b). Zhang et al. (2017) adopted this model for SGR with very low permeability and developed the LGM as an empirical method to forecast gas production. The LGM can be represented as follows:

$$q(t) = \frac{dQ}{dt} = \frac{Knbt^{n-1}}{(a + t^{n})^2} \tag{2.19}$$

where *K* is the carrying capacity.

The main benefit of LGM is that the reserve estimate is inhibited by the parameter *K* as well as the production rate, which terminates at infinite time (Zhang et al., 2017). The main assumption in this model is that the whole reservoir can be drained by a single well over a suitably long period and requires the approximation of at least two parameters, or parameters, as per the available well information (Tan et al., 2018; Nwaobi and Anandarajah, 2018).

2.7.7. Fractional Decline Curve Model (FDC)

Zuo et al. (2016) developed a new FDC model with three fitting parameters using the general solution of the fractional diffusion equations, which is a special case of the so-called Mittag-Leffler function (Zuo et al., 2016). The model is based on the anomalous diffusion phenomena that also exhibits long-tail behaviour. The FDC can be represented as follows (Yuan et al., 2020):

$$q = mE_{\propto,1}(-\lambda t^{\alpha}) = m\sum_{k=o}^{\alpha}\frac{(-\lambda t^{\alpha}}{\Gamma(\alpha k+1)}. \quad \textbf{(2.20)}$$

where m is a coefficient corresponding to well index; $E_{\propto,1}(-\lambda t^{\alpha})$ is the Mittag-Leffler function proposed by Mittag-Leffler (1903) and α, λ are fitting parameters.

In addition, they proposed a four-step scheme according to the asymptotic properties of the Mittag-Leffler function, to quantify the three parameters. The steps are as follows (Yuan et al., 2020):

1. Data Preconditioning.
2. α, λ and m Determinations.
3. Validation of these parameters with the actual data, and
4. EUR forecast comparison.

After validating the model and comparing it to the Arps model, Zuo et al. (2016), found that it produced a much smaller error compared to the Arps model. Wang et al. (2017) found that the FDC model requires iterative programming to optimize the original parameters obtained by production, and the EUR is calculated by daily production accumulation, which makes the FDC rather complicated and inconvenient.

2.7.8. Auto Regressive Integrated Model (ARIMA)

The ARIMA processes follow a stochastic behaviour used to analyse time series (Contreras et al., 2003) and is mostly used to predict demand. The ARIMA models have also proved to be excellent short-term forecasting models for a wide variety of time series because short-term factors are expected to change slowly (Raymond, 1997). A mixed autoregressive and moving average model with both components is known as an ARIMA model.

Raymond (1997) suggested that the following two questions must be answered to identify the data series in a time series analysis:

- whether the data are random; and
- have any trends.

If a series is random, the correlation between successive values in a time series is close to zero and if the observations of time series are statistically dependent on each another, then the ARIMA is appropriate for the time series analysis (Raymond, 1997).

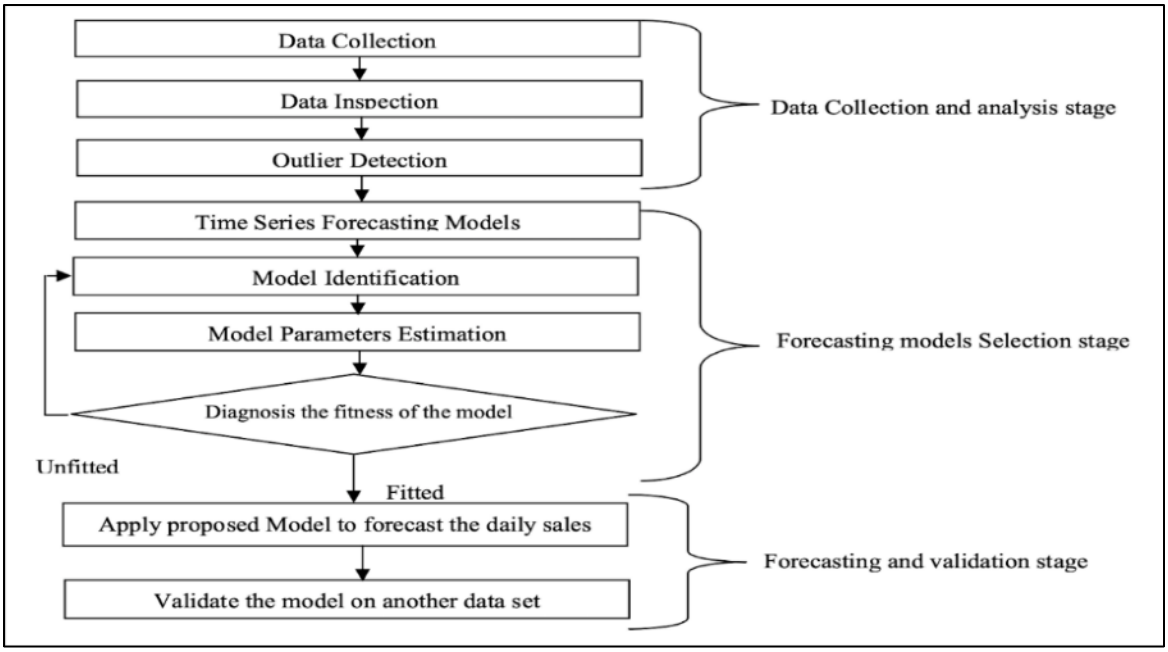


Figure 2.11: The steps for the ARIMA computation (Shukla and Jharkharia, 2011)

Box and Jenkins (Contreras et al., 2003) developed the application of the ARIMA methodology for this study of time series analysis. The Box–Jenkins methodology includes three iterative steps of model identification, parameter estimation and diagnostic checking (Zhang, 2003). This three-step model building process is typically repeated several times until a satisfactory model is finally selected and can then be used for prediction purposes (Zhang, 2003). The steps for the ARIMA model building methodology is represented in a flow-chart in **Figure 2.11**.

Contreras et al. (2003) described the steps to the ARIMA Model as follows:

- A model is identified for the observed data.
- The model parameters are estimated.
- If the hypotheses of the model are validated, go to Step 4, otherwise go to Step 1 to refine the model, and
- The model is ready for forecasting.

In an ARIMA model, the future value of a variable is assumed to be a linear function of several past observations and random errors (Contreras et al., 2003). During the past decades, researchers have been focusing more on linear models due to their simplicity in comprehension and application (Fattah et al., 2018). A disadvantage of the classical ARIMA methodology is that it requires a large number of observations to determine the best fit model for a data series (Fattah et al., 2018). An ARIMA model is labelled as an ARIMA model (p, d, q), where:

1. p is the number of autoregressive terms
2. d is the number of differences; and,
3. q is the number of moving averages.

2.7.8.1. The Autoregressive Process

This process assumes that Y_t is a linear function of the preceding values and is given by Equation (2.21).

$$Y_t = \propto_1 Y_{t-1} + \varepsilon_t \tag{2.21}$$

Generally, each observation consists of a random component i.e., a random shock, ε and a linear combination of the previous observations. $\propto_1$ in the equation is the self-regression coefficient.

2.7.8.2. The Integrated Process

The integrated processes are the archetype of non-stationary series. A differentiation of order 1 assumes that the difference between two successive values of *Y* is constant. An integrated process is defined by Equation (2.22):

$$Y_t = Y_{t-1} + \varepsilon_t \tag{2.22}$$

where the random perturbation ε_t is a white noise.

2.7.8.3. The Moving Average Process

The moving averaging process is a linear combination of the current disturbance with one or more previous perturbations. The moving average order indicates the number of previous periods embedded in the current value. Thus, a moving average is defined by Equation (2.23):

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{2.23}$$

2.7.9. Artificial Neutral Network (ANN)

Recent research activities in ANN have shown its powerful pattern classification and pattern recognition capabilities (Zhang et al., 1998). Stimulated by biological systems, particularly when made functional by research into the human brain, ANN is able to learn from and generalize from experience (Zhang et al., 1998). ANN has found increasing consideration in forecasting theory, leading to successful applications in various forecasting domains including economic, business financial and many more (Shamsuddin et al., 2008). One advantage of neural networks compared to other non-linear models is its universal model, which is capable of predicting fairly extensive functions with a high degree of accuracy. No assumptions are required for neural networks, thus neural networks conform to the characteristics of the data (Dhini et al., 2015). However, there are disadvantages, namely; in the constructing forecasting model the selection of network architecture, learning parameters and data pre-processing techniques that apply to the time series data are some of the modelling issues (Bacha and Meyer, 1992; Kaastra and Boyd, 1996).

2.7.9.1. Overview of ANN

ANN was originally developed to mimic the basic biological neural system i.e., the human brain, which is composed of a number of interconnected simple processing elements called neutrons or nodes (Zhang et al., 1998). Each node accepts an input signal which is the total “information” from other nodes, processes it locally through an activation or transfer function and produces a transformed output signal to other nodes (Zhang et al., 1998). According to Reilly and Cooper (1995), although each individual neuron implements its function rather slowly and imperfectly. Collectively, a network can perform a surprising number of tasks efficiently (Reilly and Cooper, 1995). This processing function makes ANN a dominant computational tool that is able to learn from examples and then simplify it to examples encountered before. There have been many different ANN models recommended since the 1980s. However, the most common models are the multilayer perceptions (MLP), Hopfield networks and Kohonen’s self-organising networks. For this study, the MPL will be used because it can be used in a variety of problems, especially in forecasting, because of their inherent capability random input-output mapping (Zhang et al., 1997). The MLP model consists of three interconnected layers, the input layer, the hidden layer, and the output layer. The basic unit of any ANN is the neuron or node (processor). Each node is able to sum many inputs x1, x2 ... x3 whether these inputs are from a database or from other nodes, with each input modified by an adjustable connection weight (**Figure 2.12**).

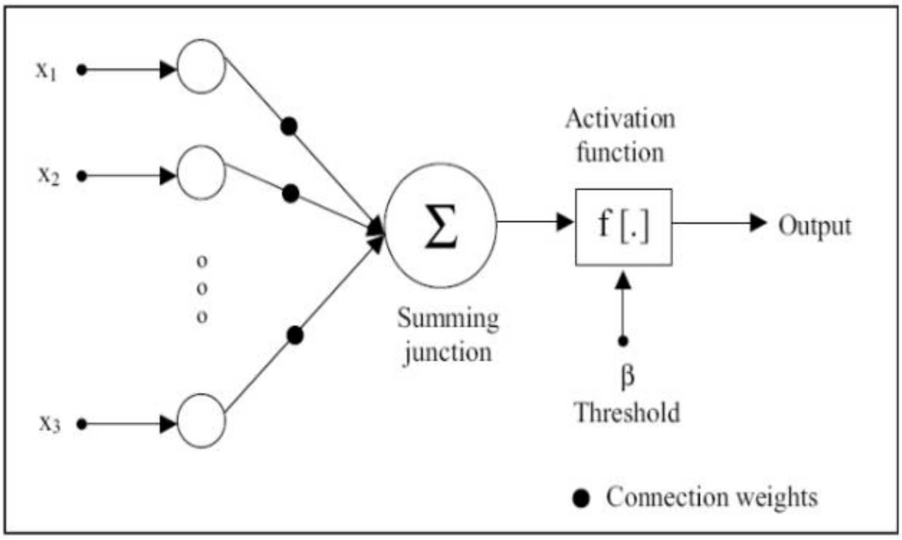


Figure 2.72: Schematic representation of the MLP ANN model (Shamsuddin et al., 2008)

The relationships that occur in the output and input layers follow Equation (2.24):

$$Y_t = \propto_0 + \sum_{j-1}^{q} \propto_j g\left(\beta_0 j + \sum_{I=1}^{p} \beta_i j Y_t - i\right) + \varepsilon_t \qquad \textbf{(2.24)}$$

where $\propto_j$ (*j* = 1,2,3, ..., q) and $\beta_i j$ (*i* = 1,2,3, ..., p; *j* = 1,2,3, ..., q) are the parameters of the model (often called the weights), *p* is the number of input points (input nodes), and *q* is the number of hidden nodes. The activation function used in the hidden layer is the logistic sigmoid function and the linear function is the output layer.

2.7.9.2. ANN Modelling Forecasting Challenges

There have been many characteristics and issues highlighted regarding the ANN model, which have to be considered. One of the key issues is the determination of the applicable architecture, as highlighted by Zhang et al. (1998), i.e., the number of layers, the number of nodes in each layer and lastly the number of arcs, which interconnect the nodes. **Table 2.2** which has been sourced from the work done by Zhang et al. (1998) summarises the literature on the modelling challenges associated with ANN.

Table 2.2: Summary of the modelling challenges associated with the ANN model taken from varies literature (Zhang et al., 1998)

Researchers	Data Type	Training/Test Size	Number input nodes	Number hidden layer: nodes	Number output nodes	Transfer fun. hidden output	Training Algorithm	Performance Measure
Chakraborty et al. (1992)	Monthly price series	90/10	8	1:8	1	Sigmoid:sigmoid	BP[4]	MSE
Cottrell et al. (1995)	Yearly sunspots	220/?	4	1:2-5	1	Sigmoid: linear	Second order	Residual variance and BIC
De Groot and Wurtz (1991)	Yearly sunspots	221/35.55	4	1:0-4	1	Tanh: Tanh	BP:BFGS LM[5] etc.	Residual variance
Foster et al. (1992)	Yearly and monthly data	*N-k/k*[6]	5,8	1:3,10	1	N/A	N/A	MdAPE and GMARE
Ginzburg et al. (1994)	Yearly sunspots	220/35	12	1:3	1	Sigmoid: linear	BP	RMSE
Gorr et al. (1994)	Student GDP	90%/10%	8	1:3	1	Sigmoid: linear	BP	ME and MAD
Grudnitski and Osburn (1993)	Monthly S, P and gold	N/A[7]	24	2:(24)(8)	1	N/A	BP	% prediction accuracy
Kang (1991)	Simulated and real series	70/24 or 40/24	4,8,2	1,2:varied	1	Sigmoid: sigmoid	GR:GR2	MSE, MAPE MAD, U-coeff.
Kohzadi et al. (1996)	Monthly cattle and wheat prices	240/25	6	1:5	1	N/A	BP	MSE, AME, MAPE

[4] Back-propagation
[5] Levenberg-Marquardt
[6] N is the number of training sample size; k is 6, 8, and 18 for yearly, monthly and quarterly data respectively
[7] Not available

Kuan and Liu (1995)	Daily exchange rates	1245/varied	varied	1:varied	1	Sigmoid: linear	Newton	RMSE
Lachtermacher and Fuller (1995)	Annual river flow and load	100%/synthetic	n/a	1:n/a	1	Sigmoid: sigmoid	BP	RMSE and Rank Sum
Nam and Schaefer (1995)	Monthly airline traffic	3,6,9 yrs./ 1 yr.	12	1:12,15,17	1	Sigmoid: sigmoid	BP	MAD
Nelson et al. (1994)	M-competition monthly	*N*-18/18	varied	1:varied	1	N/A	BP	MAPE
Schoneburg (1990)	Daily stock price	42/56	10	2(10)(10)	1	Sigmoid: sigmoid	BP	% prediction accuracy
Sharda and Patil (1992)	M-competition monthly	*N-k/k*	12 for monthly	1:12 for monthly	1,8	Sigmoid: sigmoid	BP	MAPE
Srinivasan et al. (1994)	Daily load and relevant data	84/21	14	2:(19)(6)	1	Sigmoid: linear	BP	MAPE
Tang et al. (1991)	Monthly airline and car sales	N-24/24	1,6,12,24	1:=input node #	1,6,12,24	Sigmoid:sigmoid	BP	SSE
Tang and Fishwick (1993)	M-competition	*N-k/k*	12 month 4: quarter	1:=input node #	1,6,12	Sigmoid:sigmoid	BP	MAPE
Vishwakarma (1994)	Monthly economic data	300/24	6	2:(2)(2)	1	N/A	N/A	MAPE
Weigend (1992)	Sunspots exchange rate daily	221/59 501/215	12 61	1:8,3 1:5	1 2	Sigmoid: linear Tanh: linear	BP	ARV ARV
Zhang (1994)	Chaotic time series	100 000/500	21	2:(20)(20)	1-5	Sigmoid:sigmoid	BP	RMSE

2.7.10. ANN-ARIMA Hybrid Model

The accuracy of time series forecasting is challenging for scientists (Taskaya-Temizel and Ahmad, 2005). Time series data often comprises linear as well as non-linear components (Faruk, 2010). In some cases, linear-based approaches might be more suitable than non-linear ones, due to the data characteristics. The hybrid method is a combination of ARIMA and the neural network method. According to Faruk (2010), hybrid methods have a higher degree of accuracy than neural networks. ARIMA can recognize time-series patterns well but not non-linear data patterns. On the other hand, neural networks only handle non-linear data. Therefore, hybrid models combine the advantages of ARIMA with respect to linear modelling and neural networks in terms of non-linear edge modelling (Cybenko, 1989). Notwithstanding, in some circumstances, the single model approach can outperform hybrid models ((Taskaya-Temizel and Ahmad, 2005). Mathematically, time-series data can be expressed as a combination of linear and non-linear components:

$$Y_t = L_t + N_t \tag{2.25}$$

where Y_t shows the time-series data, L_t indicates the linear components, and the non-linear components are represented by N_t. Mathematically, the neural network model for residual of *n* input nodes can be expressed as the following:

$$e_t = f(e_{t-1} + e_{t-2}, \dots, e_{t-n}) \tag{2.26}$$

where *f* is a non-linear function that is specified by the neural network. With regard to the results of the prediction error of N_t, the combination forecast using the hybrid method can be expressed as:

$$\hat{y}_t = \hat{L}_t \; + \widehat{N}_t. \tag{2.27}$$

According to Taskaya-Temizel and Ahmad (2005), two factors prevent the hybrid ARIMA–ANN method from providing good results. Firstly, the assumption of the existence of a relationship between the components of the linear and non-linear components in the data can cause performance degradation, as other model relationships (e.g., multiplicative) may exist within the data instead of linear/non-linear relationships. Secondly, no one can guarantee that the residual of the linear components will have valid non-linear patterns. Their results showed that hybrids are not always better and hence that the model selection process remains an important step, despite the popularity of hybrid models (Taskaya-Temizel and Ahmad, 2005).

Dhini et al. (2015), concurred that the hybrid method does not always give better results than the single method as the neural network method outperformed the hybrid method. Some of the possible causes for this are the basic assumptions used in the method, as well as the possibility that the residual from the linear components may not be non-linear (Dhini et al., 2015). Zhang et al. (2001) found in his work that the hybrid model was able to outperform each component model used in isolation (**Figure 2.13**).

Granger (1989) has highlighted that for a hybrid model to produce higher forecasts, the component model should be suboptimal. In general, it has been observed that it is more effective to combine individual forecasts that are based on different information sets (Granger, 1989; Perrone and Cooper, 1992). There has been limited work conducted using the hybrid model approach for shale gas prediction. Hence, it is not decisive whether this model would concur with the literature, indicating that this model would predict better results than the single models for shale gas production.

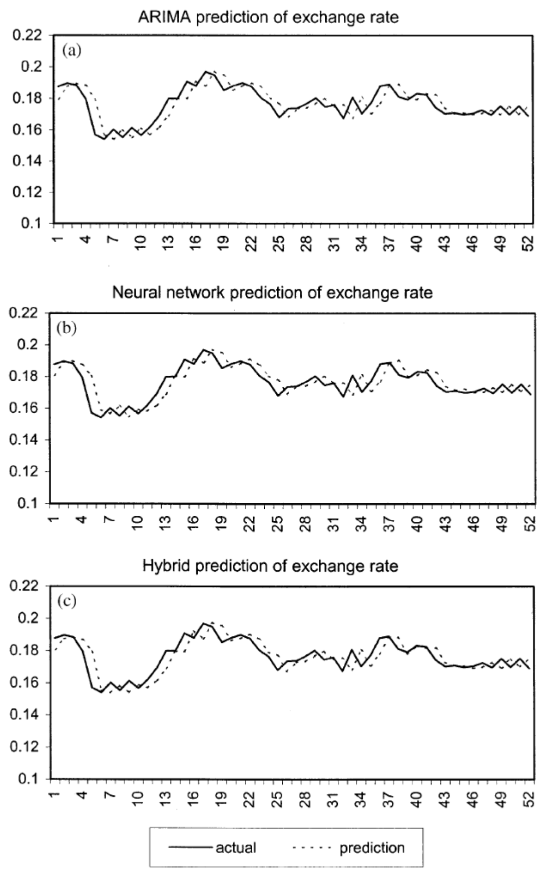


Figure 2.83: Comparison of ARIMA, ANN and Hybrid models (Zhang, 2001)

2.8. Recently Developed Decline Curve Models

2.8.1. Analytical Flow-Cell Model

Recent work by Weijermars and Nandlal (2020), has paved the way for accurate pre-drilling DCA-based well performance forecasts based on an analytical flow-cell model. The flow-cell based 2-segment DCA uses a type well to scale the production performance of newly planned wells, with the distinctive capacity to fully factor in the effects of changes in well length, fracture spacing, fracture height, fracture half-length, and variable well spacing (Weijermars and Nandlal, 2020). In summary, the 2-segment DCA can be claimed to provide the same accuracy, but is much faster and easier to use, than a gridded reservoir model (Weijermars and Nandlal, 2020).

The analytical expressions and key equations and parameters used are based on Complex Analysis Methods (CAM) that can provide solutions for fluid flow paths, velocity distribution, time-of-flight contours, and reservoir depletion patterns (Weijermars and Nandlal, 2020). The flow cell model based on CAM solves for flow interaction between multiple fractures based on material based principles (Weijermars and Nandlal, 2020). CAM-based models are gridless, have high resolution, allow fast computation of streamlines and time-of-flight contours, and have revealed the occurrence of stagnation points and flow separation surfaces between individual fractures and between adjacent well pairs (Weijermars et al., 2017a,b, 2018). The model assumes single-phase flow. The onset of well interference is determined from computing when the pressure transient of the well has reached the interwell drainage boundary (IDB) using the theoretical depth of investigation formula (2.28):

$$r_i(t) = \sqrt{\frac{k(z)t}{1688.7\phi(z)uc_t}} \qquad \textbf{(2.28)}$$

where r_i is the advances of the pressure drawdown front in ft, t is time, $\phi(z)$ is the porosity, $k(z)$ is the permeability, u is the viscosity and c_t compressibility.

The *t* in sec corresponding to the 'discovery' of a nearby well, marking the advent of True Boundary Dominated Flow (BDF), is given by (equating r_i = distance to IDB) (2.29):

$$t = \frac{r_i^2(t)1688.7\phi(z)uc_t}{k(z)} \qquad \textbf{(2.29)}$$

The timing of the onset of True BDF will vary with well spacing (distance to IDB) and the onset of BDF affects the decline of type curve wells.

2.8.2. Long short-term memory neural network (LSTM)

The LSTM neural network model is a type of RNN structure, which is widely used to solve sequence problems (Tadjer et al., 2021). The model tends to study long-term dependencies and solve the vanishing gradient problems an issue observed with the ANN model (Tadjer et al., 2021). The structure of the LSTM shown in **Figure 2.14** consists of the long term state (c_t) and three multiplicative units N with i_t, output gate (O_t) , and forget gate (f_t) and equivalently write, read, and reset information within the model's cells (Tadjer et al., 2021).

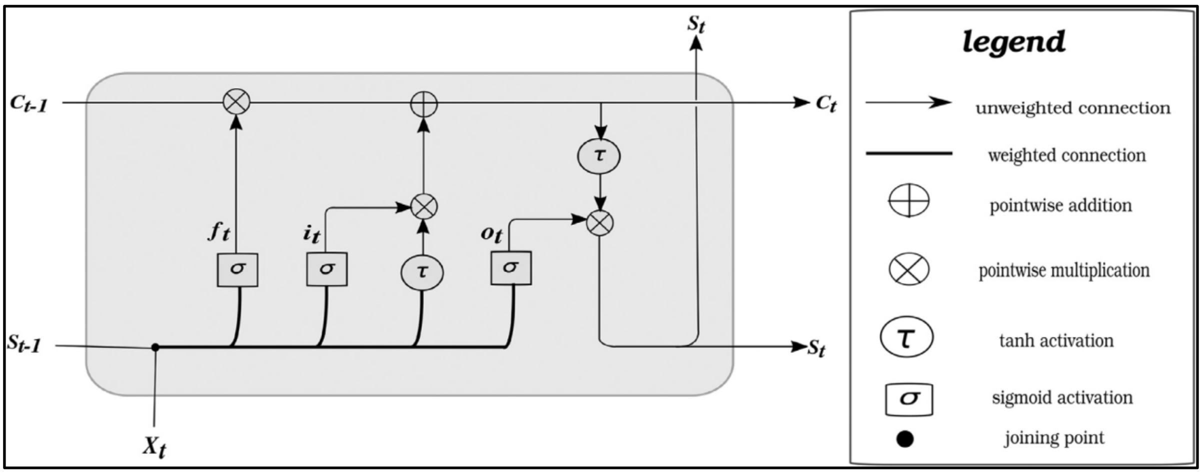


Figure 2.14: Architecture of an LSTM cell (Sagheer and Kotb, 2019)

The LSTM functions as follows (Sagheer and Kotb, 2019):

1. The first step in LSTM is to decide what information is going to be thrown away from the cell state. This decision is made by the following forget gate(f_t):

$$f_t = \sigma(X_t U^f + S_{t-1} W^f + b_f \qquad \textbf{(2.30)}$$

2. The following step is to decide which new information is going to be stored in the cell state. This step has two folds: First, the input gate (i_t) layer decides which values to be updated. Second, a tanh layer that creates a vector of new candidate values $\widetilde{C}_t$. These two folds can be described as follows:

$$i_t = \sigma(X_t U^i + S_{t-1} W^i + b_i \qquad \textbf{(2.31)}$$

$$\widetilde{C}_t = tanh(X_t U^c + S_{t-1} W^c + b_c \qquad \textbf{(2.32)}$$

3. Then, update the old cell state, C_{t-1} into the new cell state C_t , which can be given as:

$$C_t = C_{t-1} \otimes f_t \oplus i_t \otimes \widetilde{C_t} \quad \textbf{(2.33)}$$

4. Finally, decide what is going to be produced as output. This output will be based on the cell state, but will be a filtered version. In this step, the output gate o_t decides what parts of the cell state are going to be produced as output. Then, the cell state goes through tanh layer (to push the values to be between -1 and 1) and multiply it by the output gate as follows:

$$o_t = \sigma(X_t U^o + S_{t-1} W^o + b_o \quad \textbf{(2.34)}$$

$$S_t = o_t \oplus tanh(C_t) \quad \textbf{(2.35)}$$

From the previous six equations, the LSTM presents the following three groups of parameters:

1. Input weights: $\boldsymbol{U^f}$, $\boldsymbol{U^i}$, $\boldsymbol{U^o}$, $\boldsymbol{U^C}$
2. Recurrent weights: $\boldsymbol{W^f}$, $\boldsymbol{W^i}$, $\boldsymbol{W^o}$, $\boldsymbol{W^C}$
3. Bias: $\boldsymbol{b}_f$, $\boldsymbol{b_i}$, $\boldsymbol{b_o}$, $\boldsymbol{b_C}$

2.9. Economic Analysis of Decline Curve Models

Shale gas production is a profitable business for oil and gas operators, provided there is assurance that gas can be produced commercially and in a sustainable way (Guarnone et al., 2012). Historically, gas was generally only extracted from shallow wells, but recent developments in drilling technology have allowed for the lucrative extraction of natural gas from deep underground shale rock formations (Kinnaman, 2011). The present practice of shale gas economic valuation commonly uses a mean EUR and a single production decline model for the whole lease or play (Penner et al., 2013). Nevertheless, shale gas production remains economically risky because the EUR remains poorly constrained during the early stages of field development (Weijermars, 2013).

Guarnone et al. (2012) pointed out that cost estimate plays a key role because it supports the economic evaluation and support process for buying an exploration permit or not. The cost structure of a shale gas project is unlike conventional production, which makes cost estimation problematic. Guarnone et al. (2012) also mentioned appraising noteworthy OPEX-like investments involves an accurate prediction of the time when they will be required, which in turn depends on the prediction of DCA to model single wells and overall field performance. Since EUR calculated from DCA models plays a significant role in the economic analysis of shale gas development, the accuracy thereof is crucial for exploration.

Lake et al. (2013), has highlighted the difference in economics of conventional natural gas production from unconventional shale gas:

- **Total production volumes** - a conventional gas well might produce 30 to 40 billion cubic feet of gas over its life whereas a shale well would produce a fraction of this amount.
- **Rate of decline in production volumes** - shale gas wells have a very steep rate of decline compared to conventional wells, especially in the initial production period.
- **Production methods** - shale gas is trapped in rock formations that must be fractured before the gas can flow. Fracturing, which involves injecting water and sand at high pressure into the formation, is expensive and may entail environmental risks.
- **Horizontal wells** - shale gas production typically uses horizontal drilling whereas conventional gas wells are drilled using vertical wells.
- **Completion after drilling** - these wells must be fractured before the viability of producing the well can be determined. This means that the decision to drill the well is tantamount to a commitment to complete the well. This is different than a conventional gas well in which a drilling log and pressure measurements can be used to estimate the volume of recoverable reserves before the well is completed. If the well is deemed uneconomic, it can be plugged and abandoned thereby avoiding completion expenses.
- **Follow on investments** - since shale formations are typically very large, the probability of success of follow on wells (in what is commonly referred to as the resource play) is much higher than in conventional wells that seek out smaller reservoirs. This means that the opportunity to make add-on investments is greater for shale gas wells than conventional wells.

- **Exploration costs** - the vastness of shale formations means that there is little discovery risk, and few wells are unproductive. These differences indicate that the economics of extracting shale gas can be quite different from that of extracting conventional gas deposits. The first five differences tend to decrease the value of shale gas in comparison with conventional formations but the last two-favour shale very much.

They indicated that the cash flow for a gas well could be calculated from the following Equation (Lake et al., 2013):

$$[(\text{Net Gas Production (Mcf)} \times \text{Price of Natural gas per Mcf}) - \text{Total Net Operating Expenses} - \text{Severance and ad Valorem} - \text{Depletion Allownance}] \times (1 - \text{Corporate income Tax Rate}) + \text{Depletion Allow.} \quad \textbf{(2.31)}$$

The net gas production is equal to the owner's proportionate working interest in the well revenues multiplied by the estimated gross annual gas production. Total net operating expenses include the owner's share of the cash expenses required to extract, transport and sell the gas production from the well. Severance and ad Valorem expenses are taxes imposed on gas producers and the depletion allowance is a non-cash expense available to gas producers to reflect the depletion of the asset represented by the well (Lake et al., 2013). The history of shale gas development is not very extensive; consequently, there are a limited number of studies on the economic evaluation of shale gas development projects (Yuan et al., 2015).

2.9.1. Conclusion

Table 2.2 provides a summary of all ten models discussed in the chapter. The table lists the name of each model, its equation, the characteristics, strength and weakness and lastly the related references. The information collated during the literature review will be used as a premise for this study.

Table 2.3: Summary of the models

Model	Equation	Production Behaviour	Strength	Weakness	Reference
Arps Hyperbolic Decline	$q = \frac{q_i}{(1 + bD_i t)^{\frac{1}{b}}}$	linear to BDF flow	reliable and simple to use	post-production overestimation	Arps (1945) Ali and Sheng, (2015) Yuhu et al. (2016) Boah et al. (2018) Paryani et al. (2018)
Modified Hyperbolic Curve (MHD)	$q = \frac{q_i}{(1 + nD_i t)^{1/n}} \quad (D > D_{lim})$ $q = q_i \, exp(-D_{lim} t) \quad (D \leq D_{lim})$	transient and BDF flow	addresses the overestimation limitation of EUR	still unable to determine D_{lim} for production data	Robertson (1988) Yuan et al. (2016)
Power Law Exponential Decline (PLE)	$q(t) = \hat{q}_i e^{[-D_\infty t - \hat{D}_i t^{\hat{n}}]}$	transient and BDF flow	developed precisely for SGR	four unknown variables to solve	Ilk et al. (2008) Kanfar et al. (2012) McNeil et al. (2009) Seshadri and Mattar (2010) Vanorsdale (2013) Hu et al. (2018) Li et al. (2018) Paryani et al. (2018)
Stretched Exponential Decline (SEDM)	$q = q_i \, ex\,p\left[-\left(\frac{t}{\tau}\right)^n\right]$	transient flow	bounded nature of EUR and straight-line behaviours of recovery potential expression	requires sufficiently long production times	Kisslinger (1993) Johnson et al. (2009) Kanfar and Wattenbarger (2009) Valkó (2009) Valkó et al. (2010) Can and Kabir (2011) Joshi (2012) Ali and Sheng (2015)

The Extended Exponential Model (EEDM)	$\frac{\ln \frac{q}{q_o}}{t} = \beta_{\mathrm{i}} + \beta_e e^{-t^n}$	transient and BDF flow	both early and late production profiles can be captured	parameter β_l has an incomplete influence on the curve fitting and is therefore fixed	Zhou and Selim (2003) Zhang et al. (2016) Li et al. (2018) Paryani et al. (2018)
Duong's Decline	$t(a,m) = t^{-m} exp(\frac{a}{1-m}(t^{1-m} - 1))$	linear or near-linear flow	appears to fit field data from various shale plays	extended periods, a proper rate initialization against pressure is required, and in the event of water breakthrough, a and m increases	Fetkovich (1980) Duong (2011) Vanorsdale (2013) Hu et al. (2018)) Paryani et al. (2018)
Logistic Growth (LGM)	$q(t) = \frac{dQ}{dt} = \frac{Knbt^{n-1}}{(a + t^n)^2}$	long transient boundary-dominated	reserve estimate is inhibited by K as well as the production rate, which terminates at infinite time	growth is only possible up to a certain size	Clark (2011a) Clark et al. (2011b) Duong (2011) Tsoularis and Wallace (2002) Zhang et al. (2017) Hu et al. (2018) Lee and Kim (2016) Li et al. (2018)
Fractional Decline Curve Model (FDC)	$q = mE_{\propto,1}(-\lambda t^{\alpha}) = m \sum_{k=o}^{\alpha} \frac{(-\lambda t^{\alpha}}{\Gamma(\alpha k + 1)}$	anomalous diffusion phenomena	smaller error when compared to the Arps model	complicated and inconvenient	Zuo et al. (2016) Wang et al. (2017) Tan et al. (2018) Yuan et al. (2020)
Auto Regressive Integrated Model (ARIMA)	$Y_t = \propto_1 Y_{t-1} + \varepsilon_t$ $Y_t = Y_{t-1} + \varepsilon_t$ $Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$	linear	simplicity in comprehension and application	it requires a large number of observations to determine the best fit model for a data series	Cybenko (1989) Taskaya-Temizel and Ahmad (2005) Faruk (2010) Dhini et al. (2015)

Artificial Neutral Network (ANN)	$Y_t = \propto_0 + \sum_{j-1}^{q} \propto_j g\left(\beta_0 j + \sum_{I=1}^{p} \beta_i j Y_t - i\right) + \varepsilon_t$	linear and non-linear	its universal model, which is capable of predicting fairly extensive functions with a high degree of accuracy	the selection of network architecture, learning parameters and data pre-processing techniques apply to the time series data are some of the modelling issues (Refer to **Table 2.1**)	Cybenko (1989) Faruk (2010) Taskaya-Temizel and Ahmad, (2005) Dhini et al. (2015)
ANN-ARIMA Hybrid Model	$\emptyset\,(B)(1-B)^d Y_t = \theta(B)\varepsilon_t$	linear and non-linear	high degree of accuracy	approach can be found to not be fit for all types of data	Cybenko (1989) Taskaya-Temizel and Ahmad (2005) Faruk (2010) Dhini et al. (2015)
Analytical Cell Flow Model	$r_i(t) = \sqrt{\frac{k(z)t}{1688.7\phi(z)uc_t}}$ $t = \frac{r_i^2(t)1688.7\phi(z)uc_t}{k(z)}$	BDF flow	speed, affordability, simplicity, and ease of use	the flow cell model does not accurately match the actual production in certain instances	Weijermars (2017a,b, 2018) Weijermars and Nandlal (2020)
Long short-term memory neural network (LSTM)	$f_t = \sigma(X_t U^f + S_{t-1} W^f + b_f$ $i_t = \sigma(X_t U^i + S_{t-1} W^i + b_i$ $\widetilde{C_t} = tanh(X_t U^c + S_{t-1} W^c + b_c$ $C_t = C_{t-1} \otimes f_t \oplus i_t \otimes \widetilde{C_t}$ $o_t = \sigma(X_t U^o + S_{t-1} W^o + b_o$ $S_t = o_t \oplus tanh(C_t)$	major flow regimes– transient and semi-steady state flow regimes	speed up manual DCA to perform long term forecast	may suffer significant errors when used for long-term forecasts and limited interpretability	Sagheer and Kotb (2019|) Tadjer et al. (2021)

Chapter 3

Research Methodology

3.1. Introduction

The following chapter will discuss the methodology used in this research. **Figure 3.1** summarises the research methodology used, and each element will be discussed separately in this section.

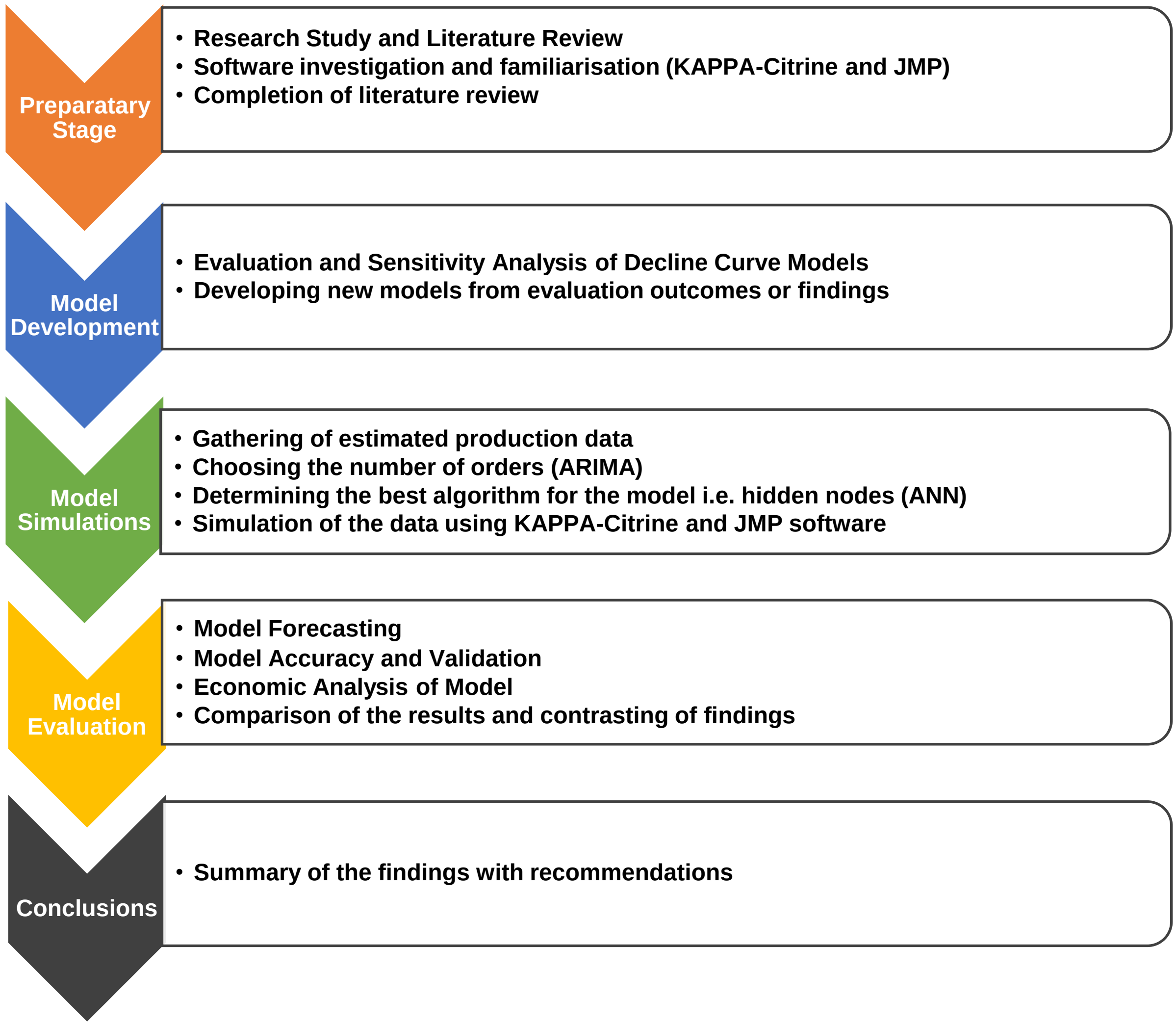


Figure 3.1: The framework for the research methodology

3.2. Preparatory Stage

The preparatory stage involved an extensive understanding of decline models through deep review of the literature on shale gas and the various published decline curve models (Chapter 2). The understanding of production and decline behaviour enables the development of models' framework and their evaluations, which assist in reviewing the model assumptions and conditions. The evaluation and sensitivity analysis of the decline curve models (Chapter 4) outcome provided guidance for the model's development. The proposed models process also involved investigating suitable software to be used in the research. There have been a number of different types of software identified or used for decline curve analysis e.g., Harmony, Serafim, TIBCO Spotfire and KAPPA- Citrine. After evaluating the pros and cons of the various software types, KAPPA-Citrine was chosen for this study.

The reasons are as follows:

- It provides two workflow methods for analysing the performance of a field.
- The entire common empirical decline models are available for fitting a decline curve to the well data.
- The regression tools can be quickly applied to the decline curves to all the wells in the field, and
- Lastly, these methods are simple to use and provide quick forecasts.

The second software types chosen was JMP software that is used primarily for time series analysis. The use of this software was predominantly chosen because JMP is simplistic and flexible enough to use, since it allows for quick interchange between models and allows one to identify what is happening in the process. Secondly, it was chosen due to the effortlessness in obtaining the software for a trial period.

3.3. Model Development

The literature evaluation process identified the limitations to the existing decline curve models and formed the premise to develop new models. The findings were as follows:

- Arps hyperbolic decline and Duong's models provided the best fit with production data.
- However, contrary to the reviewed studies when estimated production data was used in the evaluation process for the basis of this paper, using the goodness-of-fit technique, the PLE and Duong's decline models aligned the best with the production data, compared to the other models.

Based on the above findings from the evaluation process the concept of “hybrid”, which is the approach of combining models. It has been found that the hybrid model provides higher levels of accuracy compared to individual models since it combines the strengths of the different models. Therefore, the following models were developed from the hybrid model concept (Chapter 5):

- Arps-Duong’s-PLE hybrid decline model.
- Arps-Duong’s hybrid decline model.
- Arps-PLE hybrid decline model, and lastly
- Duong’s-PLE hybrid decline model.

3.4. Model Simulations

3.4.1. Shale Gas Production Data and Shale Play Parameters

A secondary data collection method was used for this study. The rationale behind the use of the secondary approach was that no shale gas production data is available because of no exploration taking place in South Africa. The secondary data would also provide for easy forecasting. However, When using this approach, one must be careful not to choose unsuitable or inadequate data in the context of the problem, for this study (Kothari, 2004).

The data was extracted from the studies of Paryani et al. (2018), Adekoya (2009) and Tan et al. (2018). They obtained the data from the Canon Shale Well (Eagle Ford), Marcellus Shale Well and Barnett Shale Well, respectively[8]. **Table 3.1** summarises the parameters for the different shale plays (Elsaig, 2016; Dong et al., 2013). The data from the table was used in the KAPPA-Citrine software to model the different decline curve models.

[8] Raw data can be found in Appendix A – A.1. Secondary Production Data

Table 3.1: Summary of shale play parameters

Reservoir	Area	Estimated Basin Area, square miles	Lateral length, ft	Net reservoir thickness, ft	Gas Content, scf/ton
Eagle Ford (Canon Shale Well)	South Texas	12000	4000	250	7-120
Marcellus Shale Well	Pennsylvania	95000	7000	50-200	60-100
Barnett Shale Well	Fort Worth Basin	5000	4800	100-600	300-350

3.4.2. KAPPA-Citrine and JMP Software

Figures 3.2 and **3.3** provides a summary of the steps followed during the simulation process using KAPPA-Citrine and JMP software.

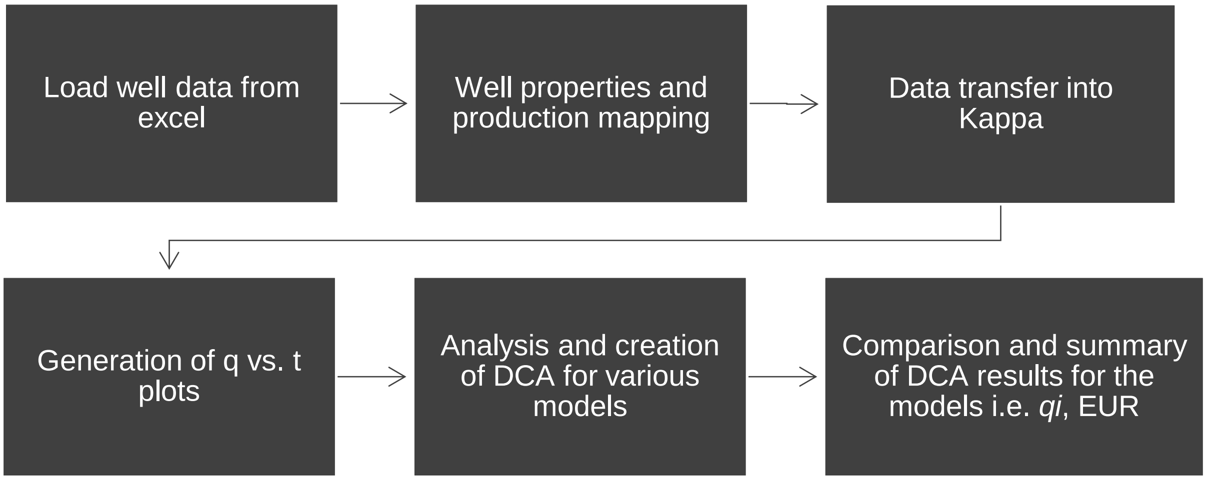


Figure 3.2: Steps for the simulation in KAPPA-Citrine[9]

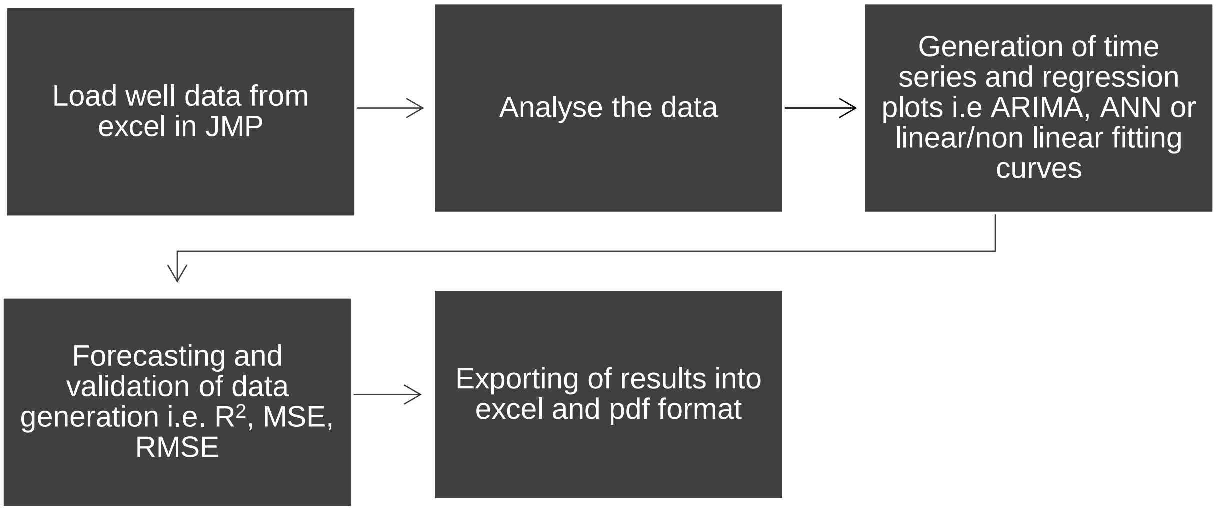


Figure 3.3: Steps for the simulation in JMP

3.5. Model Evaluation

3.5.1. Model Forecasting, Accuracy and Validation

The modelling forecasting (Chapter 7) involved three types of analysis i.e., time series, multiple regression and neural network analysis. The analysis was conducted within the JMP software. **Table 3.2** distinguishes which analysis was used for the different models.

[9] Results obtained from KAPPA-Citrine can be found in Appendix A – A.2. KAPPA-Citrine

Table 3.2: Analysis used for the different models

Multiple Regression	Time Series	Neural Network
Arps Model	ARIMA	ANN
Duong's Model	ANN-ARIMA Hybrid Model	
PLE Model		
Arps-Duong's-PLE Hybrid Model		
Arps-Duong's Hybrid Model		
Arps-PLE Hybrid Model		
Duong's-PLE Hybrid Model		

During the analysis of the data, three critical parameters such as the R^2, MSE and RMSE values were used for the accuracy assessment and validation of the models. The detailed evaluation is presented in Chapter 7.

3.5.2. Economic Analysis

The approach to the financial or economic analysis was as follows (Chapter 8):

- The data from the prediction, accuracy and validation of the forecasting model, was done in Chapter 7 to identify the model to be used in the evaluation.
- The second step was to model the EUR from the simulated or predicted data obtained from the model, which produced the results closest to the actual values, and in this study; i.e., the ANN model, this also followed a multiple regression analysis method.
- Thirdly, the results obtained from step two were used to calculate the cumulative production for the respective shale plays i.e., EUR and,
- Lastly, the EUR was used for the financial calculation.

3.6. Conclusion

The chapter covered the research methodology followed during this study. A secondary data collection method was employed to obtain production data for the Canon Shale Well, Marcellus and Barnett Shale Well. Investigation into suitable software to be used for the simulation process identified KAPPA-Citrine and JMP due to their flexibility in operation and ease of use for the operator. The steps for the software were also highlighted. The forecasting, accuracy and validation process involved a multiple regression, time series and neural network analysis. The R^2, MSE and RMSE formed critical parameters in identifying the most suitable model. The economic analysis also followed a multiple regression analysis to determine the EUR, which is a crucial parameter in the financial modelling process. All chapters provide a detailed explanation of the findings that is summarised last, with recommendations in the final chapter of the thesis.

Chapter 4

The Evaluation and Sensitivity of Decline Curve Models

4.1. Introduction

Consistently forecasting the long-term production performance of shale (unconventional) reservoirs has been a challenge (Zhang et al., 2016). The petroleum industry requires simple, useful, and speedy means of predicting production and assessing reserves. Hence, DCA has been an attractive alternative in contrast to other methods (Zhang et al., 2016). Due to the relative ease of DCA, it is considered the most used method in the industry (Zhang et al., 2016). The current DCA models will be evaluated based on their characteristics, strengths, weaknesses, and sensitivity to production data. The estimated production data was extracted from the work of Ali and Sheng (2015); Brantson et al., 2019; Paryani et al. (2018) and Tan et al. (2018).

4.2. Evaluation in the Sensitivity to Production Data[10]

4.2.1. Arps Decline Curve and the Modified Hyperbolic Decline Model

To test the behavior of the Arps hyperbolic model and the modified version shown in **Figure 4.1**, a semi log plot (log *q* versus *t*) illustrates the sensitivity of the models to various estimated field data.

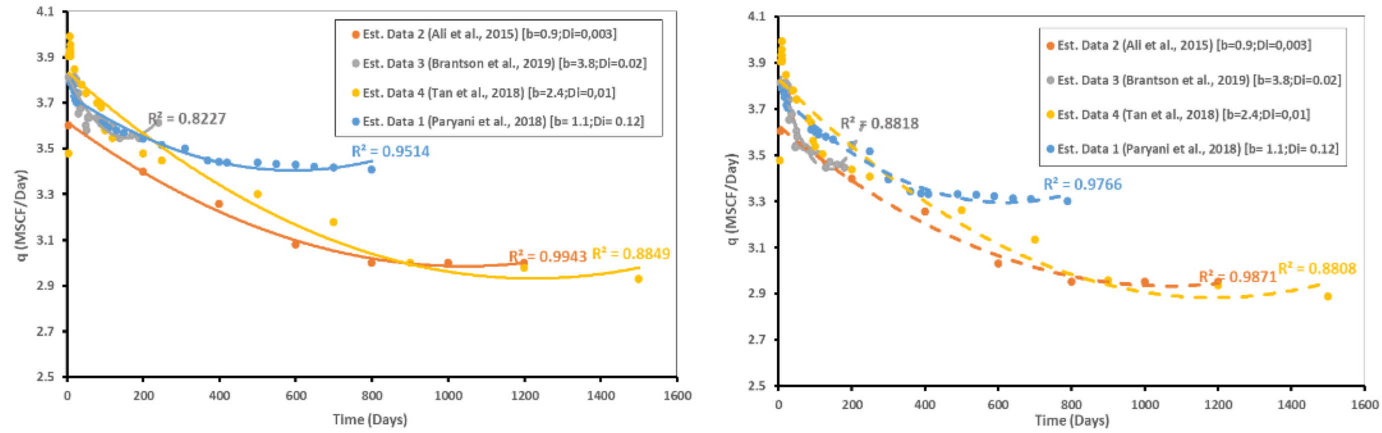


Figure 4.1: Data sensitivity using the Arps hyperbolic mode

[10] Raw data can be found in Appendix A

The R^2 values denote the goodness of fit or the degree of linear correlation, which is a measure of the level of association of a group of actual observations to the model's forecasts (Hu et al., 2018). As observed from the regression lines for the various data, the resulting fit appears to capture the trend in the data well. Arps fits Data 1 and 2 fairly, similarly for the MHD. However, the methods match the other cases poorly because they cannot model multiple flow regimes. In the case of the MHD model, there is a shift in the curves downward, which results in a change in the R^2 value.

Upon closer inspection of the EUR values for both models, which are shown in **Table 4.1**, it is evident that the MHD model corrects for the overestimation of the Arps model.

Table 4.1: Summary of the Estimated Ultimate Recovery (EUR) for Arps and the Modified Hyperbolic Decline model (MHD)[11]

EUR (bcf)	Data 1	Data 2	Data 3	Data 4
Arps Model	0.31	20.52	18.13	5.21
MHD	0.18	4.13	13.18	4.18

4.2.2. Power Law Exponential Model

The PLE model (**Figure 4.2**) uses a log–log plot (log *q* versus log *t*) to test the sensitivity of the data. The resulting fit appears to capture the trend in the data better, compared to the Arps Hyperbolic Model.

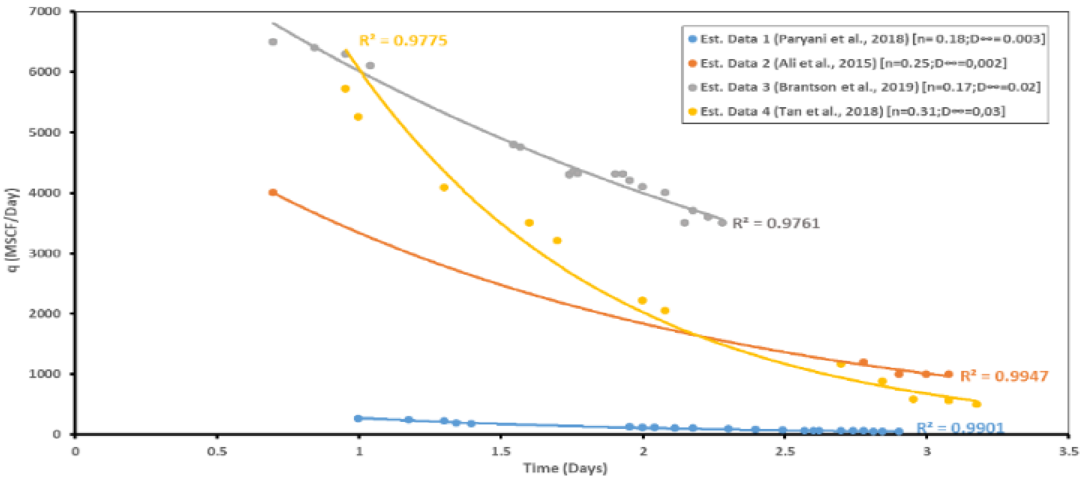


Figure 4.2: Data sensitivity using the Power Law Exponential model

[11] Results obtained from KAPPA-Citrine
Data 1 - Paryani et al., 2018
Data 2 - Ali et al., 2015
Data 3 - Brantson et al., 2019
Data 4 - Tan et al., 2018

This model fits Data 1, 2, 3, and 4 accurately. This can be attributed to the PLE model, matching production data in both transient and boundary-dominated regions.

4.2.3. Stretched Exponential Decline Model (SEDM)

Testing the behavior of the SEDM, **Figure 4.3**, which is a plot of production rate versus the cumulative production (*q* versus *Q*) to test the sensitivity of the data, the resulting fit appears to capture the trend in the data poorly. The SEDM method fits all cases inaccurately (lower R^2 values). This is due to the SEDM model's transient flow rather than boundary-dominated flow and requirement for a sufficiently long production time (usually >36 months), to accurately estimate the parameters τ and *n* (Can and Kabir, 2011).

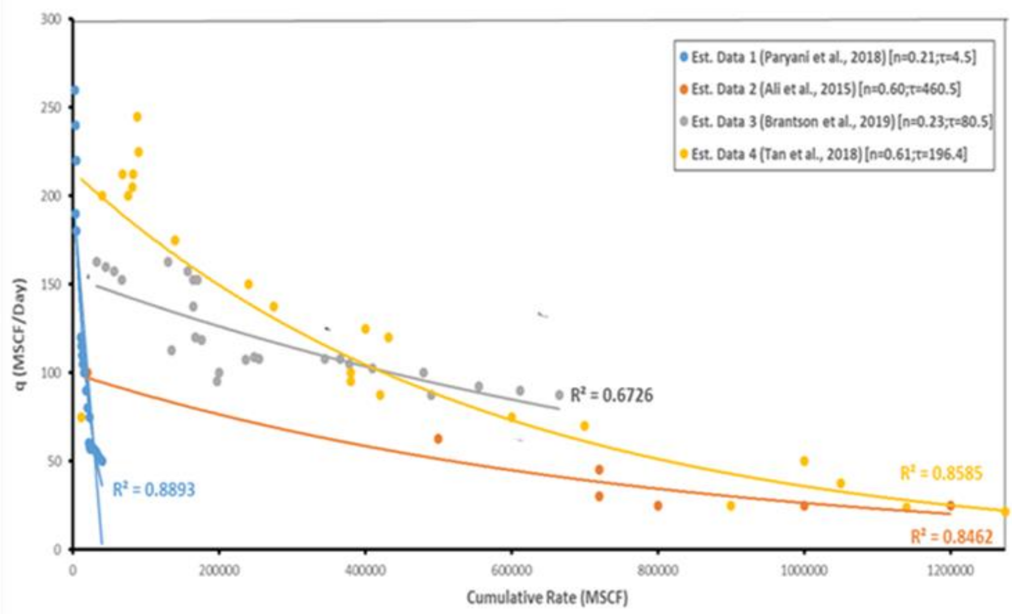


Figure 4.2: Data sensitivity using the Stretched Exponential Decline model

4.2.4. The Extended Exponential Model (EEDM)

Using the EEDM (**Figure 4.4**), which is a plot of $-ln\frac{q}{qo}/t$ versus t to test the sensitivity of the data and the resulting fit appears to also capture the trend in the data poorly. The method fits all cases inaccurately (lower R^2 values). This type of method is the best way to forecast short-term trends in the absence of recurring variations. Hence, the EEDM would only be accurate when a realistic amount of stability between the past and future is assumed.

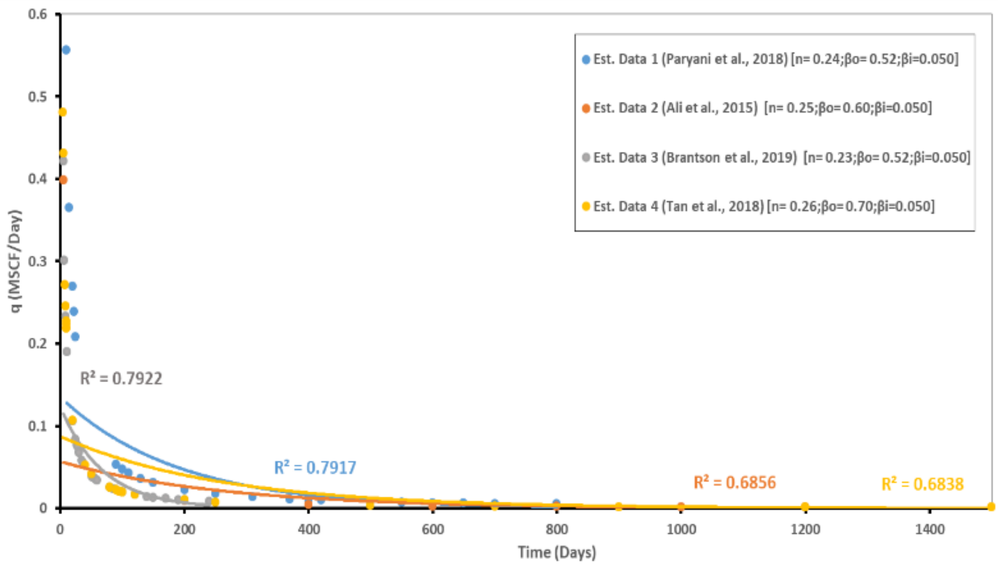


Figure 4.3: Data sensitivity using the Extended Exponential model

4.2.5. Duong's Decline Model

With the Duong's model (**Figure 4.5**), which uses a log–log linear plot (log *q* versus log *t*) to test the sensitivity of the data, the resulting fit appears to capture the trend in the data well. The method fits Data 1, 2, and 4 fairly accurately. For Case 3, the method fits the data poorly with a lower R^2 value of 0.8371. The model probably provides a good fit because it was specifically developed for unconventional reservoirs with very low permeability.

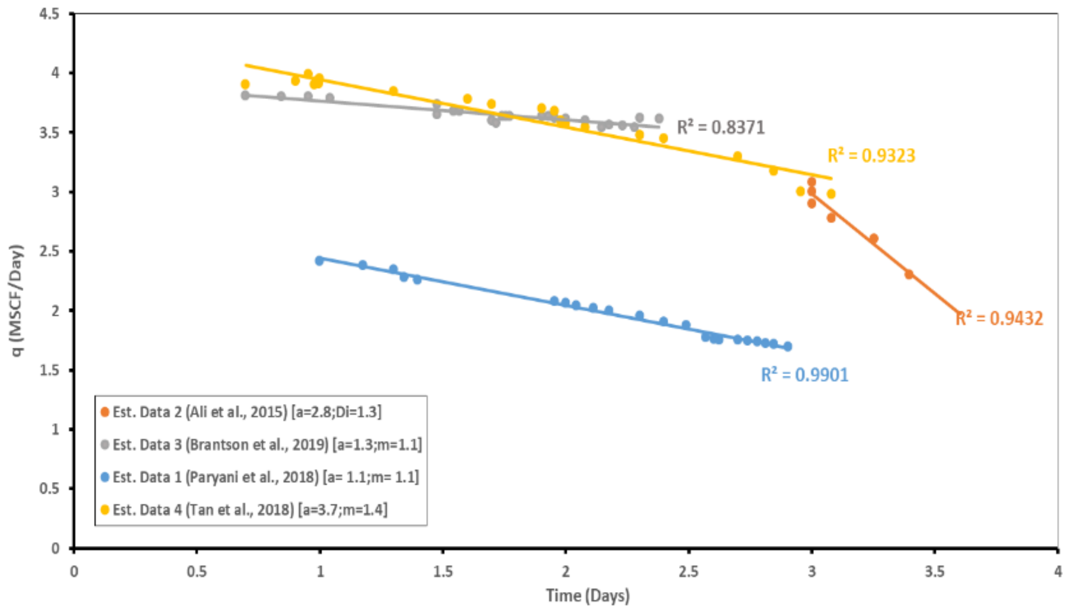


Figure 4.4: Data sensitivity using Duong's Decline model

4.2.6. Logistic Growth Model (LGM)

Figure 4.6, a plot of production rate versus time (q versus t), illustrates the sensitivity of the model to various estimated field data. As observed from the regression lines for the various data, the resulting fit appears to capture the trend in the data well. The LGM fits Data 1 and 2 fairly. However, the method matches the other cases poorly, as indicated by the lower R^2 values. This could be attributed to the data size, which is too small to yield an accurate fit, since the underlying principle of this model is population growth, which stipulates that growth is only possible up to a certain size.

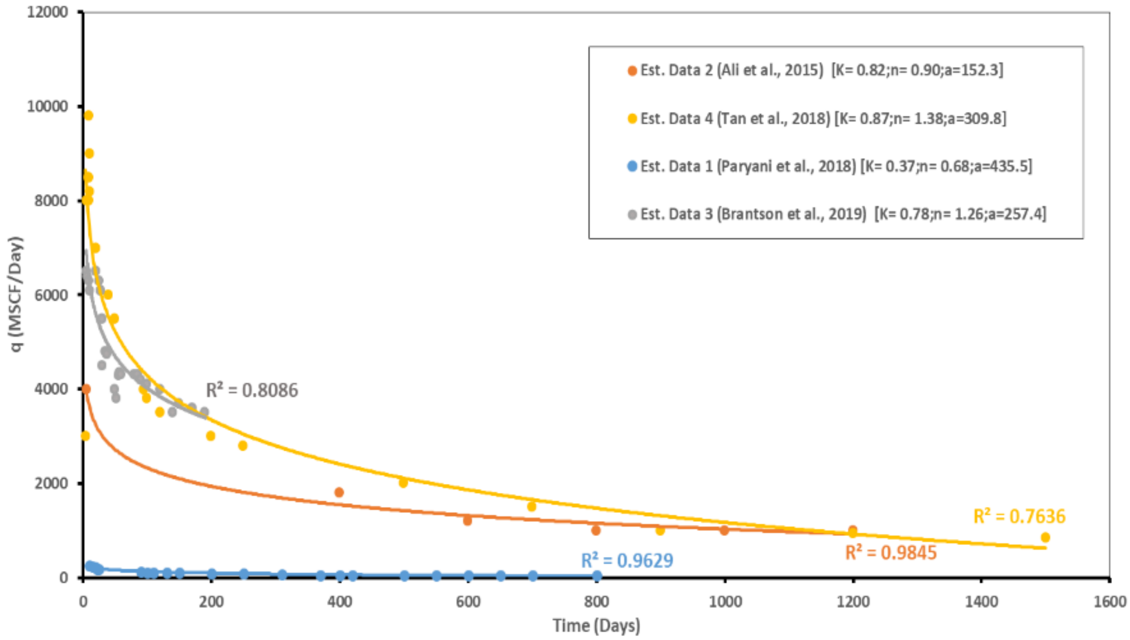


Figure 4.5: Data sensitivity using the Logistic Growth model

4.3. Accuracy of Current Decline Curve Models with Field Data

Yuhu et al. (2016) discussed comparisons of EURs with five types of decline models from single-well production data. They explained that according to the prediction results, the highest predicted EUR was gained by the hyperbolic decline model, followed by the MHD, Duong's Model, PLE and, lastly, the EEDM. Hu et al. (2018) conferred production data for wells with a production time greater than 10 years, for which the PLE decline model was recommended for multiple flows. It was also pointed out that the hyperbolic decline model predicted higher estimates of reserves than the PLE decline model. Another study that they reviewed recommended the MHD rather than the PLE decline model, which in their view was complicated.

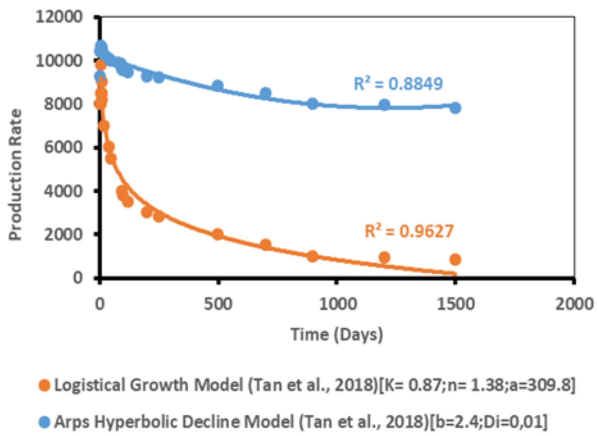
It is noted that the differences in EURs with different decline models decreases with an increase in production time (Wachtmeister et al., 2017). On the other hand, prediction consistency increases with an increase in production time. Based on the distinctive production data, the order of predicted EURs from high to low was through the hyperbolic decline model, the MHD, the PLE decline model, and the EEDM, respectively (Wachtmeister et al., 2017). The predicted EURs decreased with an increase of production time for the hyperbolic decline and the modified hyperbolic decline model. The predicted EURs increase with an increase of production time for the PLE decline and the EEDM model (Wachtmeister et al., 2017). Currently, the applicability of these different decline models is uncertain. The general trend found in their paper was that the hyperbolic decline model overestimates the production and the other decline models will still have to be investigated for reliability and accuracy (Wachtmeister et al., 2017).

In their study, Guo et al. (2017) investigated shale gas wells in the Barnett Shale Well play, where they found that from the results of goodness of fit, the hyperbolic curve fits well for both the aggregate and individual shale gas wells. On the other hand, Kenomore et al. (2018) in their production decline study of the Barnett Shale Well found that either the Arps hyperbolic or Duong's model can be used only if the historical data exceeds 10 months. They used root mean square error (RMSE) analysis and the results indicated that the Arps hyperbolic model showed better forecasting compared to the Duong's model for the top three longest production histories. Zhang et al. (2017) concurred with the findings of the Duong's model in their paper, noting that it is more accurate for linear flow and bilinear flow. However, if the production history is shorter than 18 months, this model provides unreliable results for EUR. In most circumstances, the Duong's model overestimates the total EUR. Harris (2013), in his research study of the Elm Coulee field production data, found that the Duong's method would produce the most optimistic forecasts, followed by the Arps model with a 5% minimum decline, and then the SEPD model. Shah (2013) developed new methods of combining the SEPD and Arps hyperbolic equation. The Duong's with the Arps hyperbolic equation, and the Arps super hyperbolic, combined with the Arps hyperbolic decline equation. The author found that the SEPD and Arps hyperbolic equation gave the most conservative results of all the methods in this study, even if there was insufficient data available. This equation can also work without enough BDF data being available.

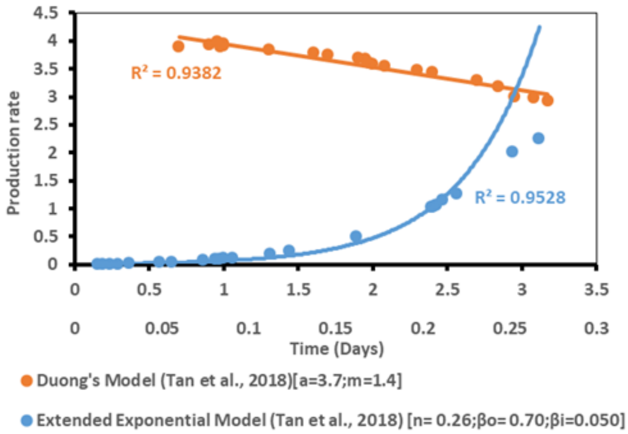
Hu et al. (2018) studied DCA techniques for the Eagle Ford and Austin Chalk reservoirs. They found that in the case of the Eagle Ford reservoir, the MHD and the Duong's model provided the highest EUR estimations and the two lowest matching errors, while the PLE decline model with $D_\infty \neq 0$ produced the lowest EUR estimates with the highest matching errors in all cases. In another study, according to the results of goodness of fit (R^2 and N-RMSE), the hyperbolic model fits well with aggregated well data and with individual wells (Zhang et al., 2017). Furthermore, in their study, Hu et al. (2018) explained that the LGM and PLE model with $D_\infty = 0$ gave production projections that were neither too positive nor too traditional, with modest matching errors. Therefore, they recommend both the MHD and Duong's model for this reservoir. However, Zhang et al. (2016) developed the EEDM and verified their model using field data from the Eagle Ford. They found this model to be more rigorous in that it included the effects of interference among adjacent fractures, variable permeability, and discontinuous pressure distribution, all of which are difficult to capture and model with other DCA methods (Zhang et al., 2016). In the case of the Austin Chalk reservoirs, all DCA methods resulted in similar EUR forecasts and matching errors. Hence, any method can be used (Vanorsdale, 2013).

Figure 4.7, which uses estimated production data versus time values, indicates that when using the R^2 values as a goodness of fit to determine the accuracy of the different decline models, the SEDM, followed by the LGM, EEDM, PLE, Duong's decline model and, lastly, the hyperbolic decline model, would predict the EUR accurately.

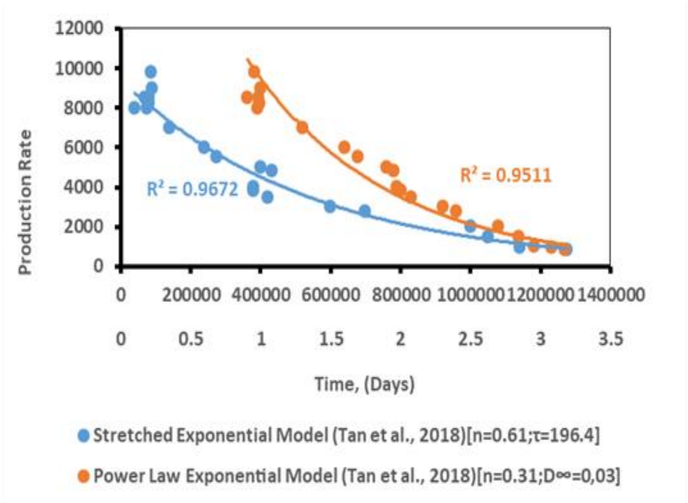
During their case study analysis, Paryani et al. (2018) found that the LGM, PLE, and Duong's models overcame Arps limitations to a certain degree. The PLE model always predicted the lowest forecasts of all the models with the most conservative production forecasting and reserve estimation. Duong's model performed the best when less noisy production data was available. However, erratic EUR was observed, which indicates that this model required further improvements (Paryani et al., 2018). The LGM gave reasonable EUR estimates when compared to the Arps model. There was an 81% fit of the wells' past production rate and cumulative production. The LGM also appears most effective at historically matching past production and predicting finite reasonable EUR. However, Tan et al. (2018) found that due to the constraints of *K* and the vanishing production rate at infinity time, the LGM provides a finite estimate of EUR. They also concluded by using normalized and logarithmic rate-time residuals that the limitations of the Arps model can be overcome and accuracy can be improved in cases of unconventional reservoirs.



(a)



(b)



(c)

Figure 4.7: Estimated production data to determine goodness of fit for the accuracy of the different decline models (a) Duong's model vs. EEDM, (b) LGM vs. Arps Hyperbolic Model and (c) SEDM vs. PLE

4.4. Conclusion

Shale gas reservoirs have become an essential source for providing natural gas globally; and the process of hydraulic fracking has been used in the extraction of shale gas. During the fracking process, there are different flow regimes, which occur during the life cycle of SGRs being fracture linear flow, fracture boundary flow, matrix linear flow, and matrix boundary flow. They are significant because they impact both the production and declining behavior of SGRs.

Based on previous studies, it was found that the Arps hyperbolic decline, the MHD and Duong's models provided the best fit with production data. However, contrary to the reviewed studies when estimated production data was used in the evaluation process for the basis of this paper, using the goodness-of-fit technique, the PLE and Duong's decline models aligned with the production data, compared to the other models.

It is evident from the accuracy assessment that decline curve modelling impacts the EUR of SGRs, and it was observed that all decline models yield a different EUR result, which is either over or under-estimated. Studies have revealed that the production time significantly impacts the EUR, depending on which decline model is being used. When each model was assessed for accuracy; once again using the goodness-of-fit technique, the results indicated that the SEDM, followed by the LGM, EEDM, PLE, Duong's decline model and, lastly, the hyperbolic decline model, align with the production data.

It is evident from the decline curve evaluation that there are advantages to using the current DCA models. However, they also have limitations associated with them, which have to be addressed.

Chapter 5

Model Development

5.1. Introduction

It is evident from the Evaluation and Sensitivity of Decline Curve Models (Chapter 4) that there are advantages associated with using the current DCA models. However, they do come with limitations, which need to be addressed. The Arps model is inaccurate within the TFR, and the Duong's model is inaccurate within the BDF. Although the PLE model incorporates both these flow regimes and was specifically developed for SGRs, the model has its own shortcomings. Hence, there is scope to develop a new decline model or a new method to predict the recovery of SGRs more accurately. The new method proposed is to combine the above-mentioned methods i.e., to evaluate hybrid models. As the PLE and Duong models the transient flow well and Arps is widely used for BDF, the new method combines the methods to achieve the required objectives and eliminates the shortcomings of the stand-alone models.

For the Arps model where the values are within $0 \leq b \leq 1$ represents a hyperbolic decline. The Arps approach usually limits the value of b to ($0 \leq b \leq 1$) (Shah, 2013). Lee and Sidle (2010) showed that $b > 1$ gives physically impossible results when Arps cumulative production equation is evaluated at infinite time. Ilk et al. (2008) introduced the PLE decline method to better fit and forecast tight gas and shale production. The PLE models the loss-ratio uniquely by assuming that the loss-ratio follows a power law function at early time and becomes constant at late time. The Duong equation models transient flow, so it assumes prolonged production within this flow regime (Shah, 2013). Typical ranges for the Duong's parameters are $1 < m < 2$ and $1 < a < 2$ (Paryani, et al., 2017).

Thus, there are four new hybrid models being proposed for this study: -

- Arps-Duong's-PLE hybrid decline model.
- Arps-Duong hybrid decline model.
- Arps-PLE hybrid decline model and lastly, and
- Duong's-PLE hybrid decline model.

5.2. The Arps-Duong's-PLE Hybrid Decline Model

The first proposed method incorporates the three models Arps, Duong's and PLE decline models. The Arps model only considers BDF while Duong's and PLE models consider TFR. The PLE model also considers BDF and has been specifically developed for SGRs. Hence, by combining the three models, the limitations from each are presumed to be minimised or eliminated.

5.2.1. Assumptions and Conditions

Table 5.1 summarises the different model behaviours, assumptions and conditions taken into consideration for the combined models.

Table 5.1: Summary of the behaviour, assumptions, conditions and parameters for the Arps-Duong's-PLE hybrid decline model

	Model	Equation	Production Behaviour	Assumptions	Condition
1	Arps Hyperbolic Model	$qt = q_i(1 + bD_i t)^{-1/b}$	BDF	Decline parameter, *b*, defines the decline behaviour	$0 \leq b \leq 1$ *D* is changing
2	Duong's Model	$qt = q_i t(a.m) + qm$	TFR	Very low permeability and long periods of transient flow	$1 < m < 2$
3	Power Law Model	$qt = q_i e^{[-D_\infty t - D_i tn]}$	BDF and TFR	Approximates the rate of decline with a power law decline	D_∞ is constant at late time

5.2.2. Model Derivation

STEP 1: Simplifying Duong's Formulae

$$qt = q_i t(a.m) + qm$$

$$\frac{qt}{q_i} = t(a.m) + \frac{qm}{q_i}$$

$$t(a.m) = \frac{qm}{q_i} - \frac{qt}{q_i}$$

$$t(a.m) = \frac{qm - qt}{q_i} \tag{5.1}$$

Where $t(a.m) = t^{-m} exp\,(\frac{a}{1-m}(t^{1-m} - 1)$

STEP 2: Using Eq. (5.1) substitute into Eq. (1) from Table 5.1 and assuming *qm* = 0

$$qt = q_i(1 + bD_it)^{-1/b}$$

$$qt = q_i[1 + bD_i[\frac{qt}{q_i}]]^{-\frac{1}{b}}$$

$$qt = q_i\ [\frac{q_i + bD_i(qt)}{q_i}]^{-\frac{1}{b}}$$

$$qt = q_i{}^{(1-(-\frac{1}{b}))}[q_i + bD_i(qt)]$$

$$qt = q_i{}^{(\frac{b+1}{b})}[q_i + bD_i(qt)]$$

$$qt = \left[q_i{}^{\frac{b+1}{b}} \times q_i\right] + qt \tag{5.2}$$

STEP 3: Combining Eq. (3) from Table 5.1 and Eq. (5.2)

$$q_ie^{[-D_it-D_itn]} = \left[q_i{}^{\frac{b+1}{b}} \times q_i\right] + qt$$

$$e^{[-D_\infty t-D_itn]} = \frac{{}^{q_i}b+1}{{}_{q_i}\ b} + \frac{qt}{q_i}$$

$$In\ e^{[-D_\infty t-Ditn]} = \ln\frac{b+1}{b} + \frac{qt}{q_i}$$

$$-D_\infty t - D_itn = In\ \frac{b+1}{b} + \frac{qt}{q_i}$$

$$t\ (-D_\infty - D_in) = In\frac{b+1}{b} + \frac{qt}{q_i}$$

$$\boldsymbol{\frac{qt}{q_i} = t\ (-D_\infty - D_in) - In\frac{b+1}{b}} \tag{5.3}$$

where q_i is the initial flow rate in STB/day or Mscf/day, D_∞ is the decline rate at long-term periods while D_i is the initial decline constant, which are both measured in days $^{-1}$, n is the time exponent and b is the decline exponent.

When the three models, Arps, Duong's and PLE were combined, the results indicated that the conditions from the Arp's and PLE models were dominant. Therefore, the Arps condition of $0 \leq b \leq 1$ for a hyperbolic decline was satisfied, According to Shah (2013), the decline exponent must be within this range to apply the Arps curves correctly. Also, with the PLE model, the "loss rate" is assumed to follow a power law function initially and then it becomes constant during the later time period. Hence, it can be presumed that during the decline process there will be a switching point from a hyperbolic to an exponential decline.

5.3. The Arps-Duong Hybrid Decline Model

The second proposed model incorporates the two developed models. Arps only considers BDF while Duong's considers TFR, hence, both flow regimes will be taken into account when combining these two models.

5.3.1. Assumptions and Conditions

Table 5.2 summarises the different model behaviours, assumptions and conditions taken into consideration for the combined models.

Table 5.2: Summary of the behaviour, assumptions, conditions and parameters for the Arps-Duong hybrid decline model

	Model	Equation	Production Behaviour	Assumptions	Condition
1	Arps Hyperbolic Model	$qt = q_i(1 + bD_it)^{-1/b}$	BDF	Decline parameter, *b*, defines the decline behaviour	$0 \leq b \leq 1$ *D* is changing
2	Duong's Model	$qt = q_it(a.m) + qm$	TFR	Very low permeability and long periods of transient flow	$1 < m < 2$

5.3.2. Model Derivation

STEP 1: Simplifying Duong's Formulae

$$qt = q_it(a.m) + qm$$

$$q_i = \frac{qm - qt}{t\,(a.m)} \quad \textbf{(5.4)}$$

STEP 2: Using Eq. (5.4) substitute into Eq. (1) from Table 5.2

$$qt = q_i(1 + bD_it)^{-1/b}$$

$$\boldsymbol{qt} = \left|\frac{\boldsymbol{qm - qt}}{\boldsymbol{t}}\right| [\mathbf{1} + \boldsymbol{bD_it}]^{-\frac{1}{b}} \quad \textbf{(5.5)}$$

where q_t is the flow rate at time, *t* in STB/day or Mscf/day, D_i is the initial decline constant, in days $^{-1}$ and *b* is the decline exponent.

When the two models, Arps and Duong's were combined, the result indicated that the conditions from the Arps model were prevalent. Therefore, the Arps condition of $0 \leq b \leq 1$ for a hyperbolic decline will be observed with a changing D_i value.

5.4. The Arps-PLE Hybrid Decline Model

The third proposed model incorporates the Arps and PLE Models. These models consider BDF and TFR flows. Since the PLE model was developed specifically for SGRs it would be advantageous to evaluate these two models as combined due to both being simple equations to use.

5.4.1. Assumptions and Conditions

Table 5.3 summarises the different model behaviours, assumptions and conditions taken into consideration for the combined models.

Table 5.3: Summary of the behaviour, assumptions, conditions and parameters for the Arps-PLE hybrid decline model

	Model	Equation	Production Behaviour	Assumptions	Condition
1	Arps Hyperbolic Model	$qt = q_i(1 + bD_it)^{-1/b}$	BDF	Decline parameter, b, defines the decline behaviour	$0 \leq b \leq 1$ D is changing
2	Power Law Model	$qt = q_ie^{[-D_\infty t - D_itn]}$	BDF and TFR	Approximates the rate of decline with a power law decline	D_∞ is constant at late time

5.4.2. Model Derivation

STEP 1: Equating Eq. (1) and (2) from Table 5.4

$$q_ie^{-D_\infty t - D_itn} = q_i(1 + bD_it)^{-1/b}$$

$$q_i[-D_\infty t - D_itn] = -\frac{1}{b}\ In\ q_i(1 + bD_it)$$

$$q_it[-D_\infty - D_in] = -\frac{1}{b}\ In\ q_i(1 + bD_it)$$

$$\boldsymbol{t[-D_\infty - D_in] = \frac{-\frac{1}{b}}{qt}In(1 + bD_i)} \tag{5.6}$$

where q_t is the flow rate at time, t in STB/day or Mscf/day, D_∞ is the decline rate at long-term period while D_i is the initial decline constant, which are both measured in days $^{-1}$, n is the time exponent and b is the decline exponent.

The combined Arps-PLE hybrid decline model, the parameters of the PLE model are observant. This model introduces the n and D_∞ Arps model. The major advantage of this model is that the extra variables enable the model to account for the transient flow. Hence, the "loss rate" would assume a power law function initially and then become constant during the later time period. However, with this hybrid model the number of variables to solve is reduced to three from four, which was evident in the PLE model.

This according to Hu et al. (2018) results in several degrees of freedom and may be clumsy to use or solve. Hence, with combining the models the number of degrees of freedom is minimised.

5.5. The Combined Duong's and PLE Models

The fourth proposed model incorporates the Duong's and PLE Models. These models both consider TFR.

5.5.1. Assumptions and Conditions

Table 5.4 summarises the different model behaviours, assumptions and conditions taken into consideration for the combined models.

Table 5.4: Summary of the behaviour, assumptions, conditions and parameters for the Duong's-PLE hybrid decline model

	Model	Equation	Production Behaviour	Assumptions	Condition
1	Duong's Model	$qt = q_i t(a.m) + qm$	TFR	Very low permeability and long periods of transient flow	$m > 1$
2	Power Law Model	$qt = q_i e^{[-D_\infty t - D_i tn]}$	BDF and TFR	Approximates the rate of decline with a power law decline	D_∞ is constant at late time

5.5.2. Model Derivation

STEP 1: Using Eq. (5.4) and substituting into Eq. (2) from Table 5.4

$$qt = \frac{qt - qm}{t(a.m)} e^{[-D_\infty t - D_i tn]}$$

$$\ln qt = \frac{qt - qm}{t(a.m)} [-D_\infty t - D_i tn]$$

$$\ln qt = \frac{qt - qm}{1} \frac{1}{t(a.m)}\ t\ [-D_\infty t - D_i tn]$$

$$\ln qt = \frac{qt - qm}{1} \frac{1}{t(a.m)}\ t\ [-D_\infty t - D_i tn]$$

$$\frac{\boldsymbol{In\ qt}}{\boldsymbol{qm}} = \boldsymbol{t}(-\boldsymbol{D}_\infty - \boldsymbol{D_i n}) \quad \textbf{(5.7)}$$

where q_t is the flow rate at time, t in STB/day or Mscf/day, q_m is the intercept of the plot of qt vs. t, D_∞ is the decline rate at long-term period while D_i is the initial decline constant, which are both measured in days $^{-1}$, n is the time exponent.

In the Duong's-PLE hybrid decline model, parameters from both models contribute to the decline process. Duong's method was developed on the basis that production rate and time have a power law relation or form a straight line when plotted on a log-log scale (Duong, 2011). The production trend will deviate from a log-log straight line when BDF is reached (Kanfar and Wattenbarger, 2012). The PLE model also follows an exponential decline. The "loss rate" assumes that it will follow a power law function initially and then becomes constant during later periods.

5.6. Conclusion

During the model development, four hybrid models were developed, Arps-Duong's-PLE hybrid decline model, Arps-Duong hybrid decline model, Arps-PLE hybrid decline model and Duong's-PLE hybrid decline model. The primary objective of the new models is to eliminate the shortcomings of the stand-alone models.

In the Arps-Duong's-PLE hybrid decline model, the result indicates that the conditions from the Arp's and PLE models are dominant. Therefore, the Arps condition of $0 \leq b \leq 1$ for a hyperbolic decline is noted. Also with the PLE model, the "loss rate" assumes to follow a power law function initially, and then becomes constant during the later time period. Therefore, there will be a switching point from a hyperbolic to an exponential decline.

When the two models, Arps and Duong's are combined, the result indicates that the conditions from the Arps model is prevalent. Therefore, the Arps condition of $0 \leq b \leq 1$ for a hyperbolic decline will be observed with a changing D_i value.

In the combined Arps-PLE hybrid decline model, the parameters of the PLE model are observed. This model introduces the n and D_∞ the Arps model. The major advantage of this model is the extra variables that enable the model to account for the transient flow. The "loss rate" would assume a power law function initially and then become constant during the later time period. Nevertheless, with this hybrid model, the number of variables to solve is reduced from three to four, which was evident in the PLE model. Henceforth, with combining the models the number of degrees of freedom is minimised.

Lastly, with the Duong's-PLE hybrid decline model, parameters from both models contribute to the decline process. The PLE model also follows an exponential decline. The "loss rate" assumes following a power law function initially and then becomes constant during later periods. The next chapter will focus on the simulations of the developed hybrid decline models.

Chapter 6

Simulation Models

6.1. Introduction

Chapter 5 covered the development of the simulation models. In this section, the simulations of the models will be examined. The variable used in this investigation is flow rate, $q(t)$ in STB/day, monitored over a period of time (T) in days. The estimated data was extracted from the research conducted by Paryani et al. (2018) and Adekoya (2009).

Paryani et al. (2018) sourced the data from the Cannon Shale Well located in Karnes County while Adekoya's (2009) data was from the Marcellus Shale Well located in the Appalachian region of the United States of America. KAPPA-Citrine and JMP software was used for the simulation of the models.

6.2. Production Behaviour[12]

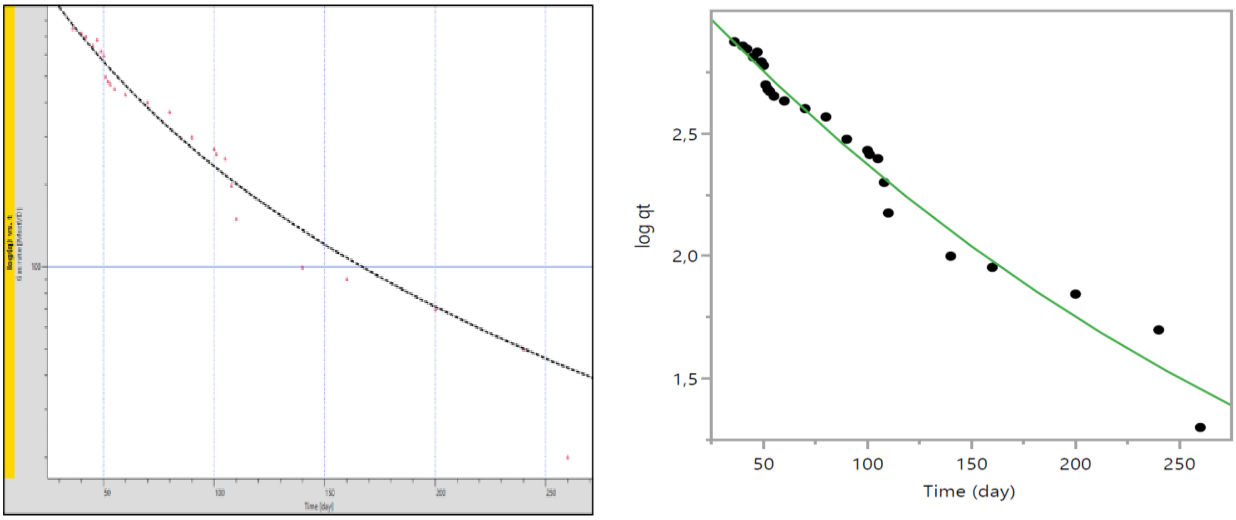
6.2.1. Arps Decline Curve Model

Arps decline curve analysis is the most commonly used method of estimating ultimate recoverable reserves and future performance (Qu and Lin, 2017). The model process is based on the following vital assumptions: that past operating conditions will remain unaffected; that a well is produced at or near capacity; and that the well's drainage remains constant and is produced at a constant bottom-hole pressure (Robertson, 1988). Notably, the Arps model is only applicable in pseudo-steady flows when the flow regime transfers from linear flows to BDF (Hu et al., 2018).

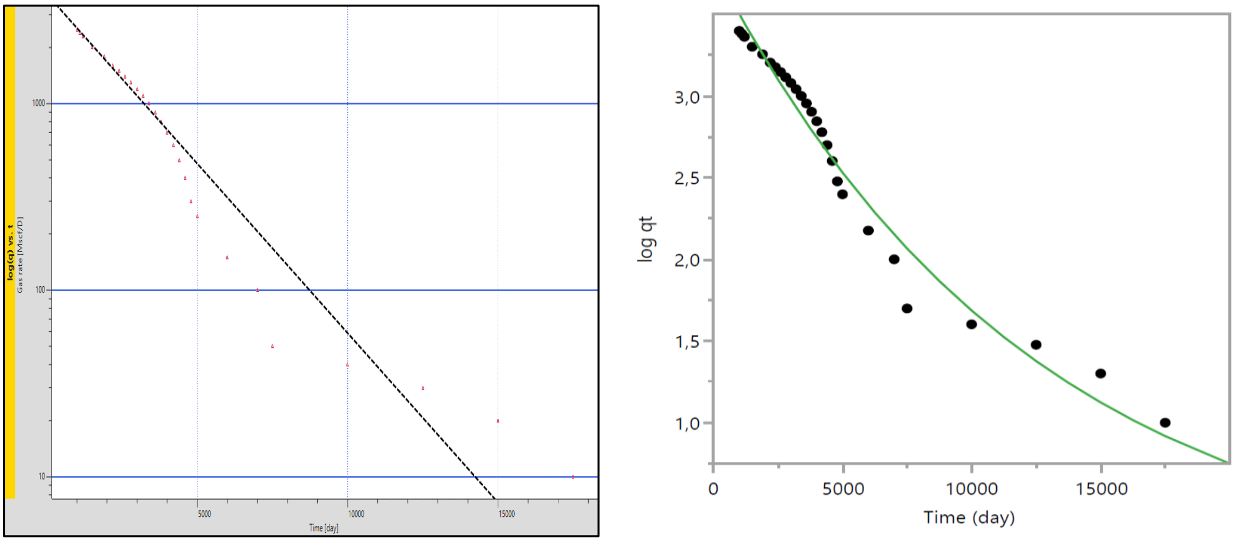
This indicates that the Arps Equations are not applicable to the production forecasting of the entire decline process of horizontal wells in low-permeability reservoirs (Bagozzi and Yi, 1988). The most commonly employed hyperbolic form of Arps decline Equation (2.2) is used for shale reservoirs.

[12] Raw data can be found in Appendix A

The hyperbolic decline equation is suitable to use due to the “best fit” that it provides for the long transient linear-flow regime observed in shale gas wells with *b* values greater than unity (Brantson et al., 2019).



(a)



(b)

Figure 6.1: KAPPA-Citrine and JMP graphs for an Arps plot for (a) Canon Shale Well and (b) Marcellus Shale Well using estimated production data

Based on **Figure 6.1**, the rate of decline is assumed to follow a logistic growth. The production data points do not all appear to fall on the fitting curve (Refer to Table A1 and A2 in Appendix A for data). This can be seen from a *log qt vs. t* plot, the R^2 values of 0.9717 (Canon Well) and 0.9692 (Marcellus Well) were observed. The summary of the Arps model simulation can be seen in **Table 6.1.**

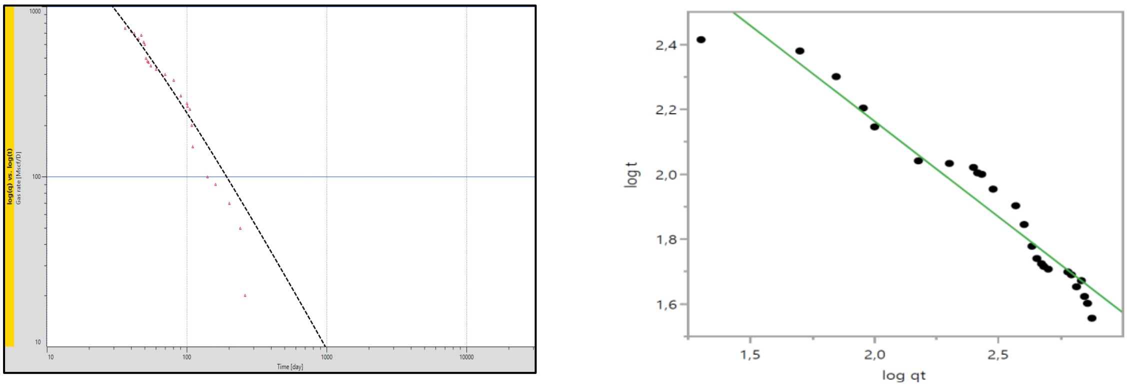
Table 6.1: Summary of the Arps decline model behaviour, assumptions, conditions and parameters

	Production Behaviour	Assumptions	Conditions	Parameters
Canon Shale Well	Boundary Dominated Flow (BDF)	Decline parameter, *b*, defines the decline behaviour The rate of decline approximated using a logistic growth decline	$0 \leq b \leq 1$	$b = 0.383$ $D_i = 3.53$ $q_i = 2183$ EUR = 0.100 $R^2 = 0.9717$ RMSE = 0.0725
Marcellus Shale Well	Boundary Dominated Flow (BDF)		$0 \leq b \leq 1$	$b = 0.00$ $D_i = 0.042$ $q_i = 3864$ EUR = 9.24 $R^2 = 0.9692$ RMSE = 0.1279

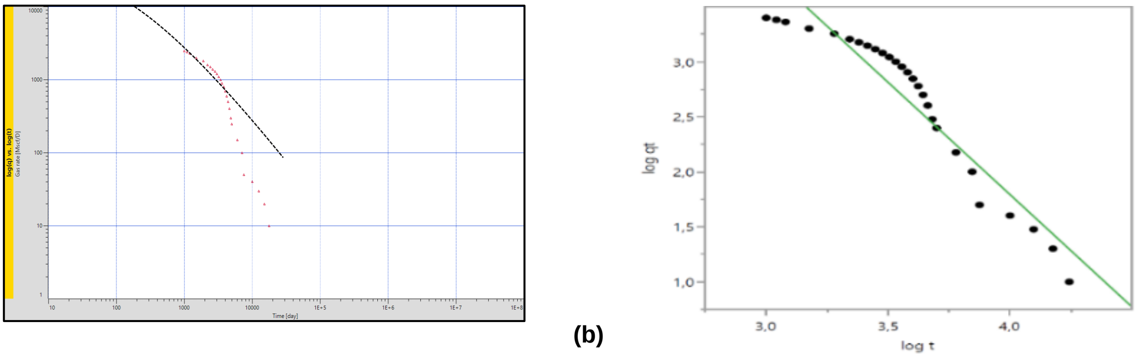
6.2.2. Duong's Decline Curve Model

Duong (2011) presented an unconventional rate decline method to evaluate the performance of shale gas wells that does not depend on the fracture types. The model assumes linear or near-linear flow, as indicated by a log–log plot of rate over cumulative production versus time, which yielded a straight-line tendency (Duong, 2011). The rate is calculated in the model using Equation (2.16).

From **Figure 6.2**, the rate of decline is assumed to follow a linear decline. As in the case of the Arps decline model, the production data points do not appear to fall on the fitting curve (Refer to Table A1 and A2 in Appendix A for data). The R^2 values for a plot of *log qt vs. log t* were 0.9402 (Canon Well) and 0.8999 (Marcellus Well), respectively. The summary of the Duong's model simulation can be seen in **Table 6.2**.



(a)



(b)

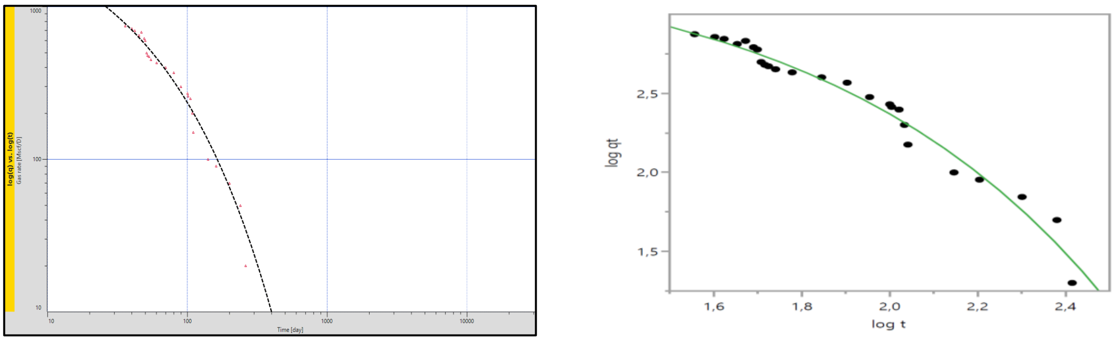
Figure 6.2: KAPPA-Citrine and JMP graphs representing the Duong's plot for (a) Canon Shale Well and (b) Marcellus Shale Well using estimated production data

Table 6.2: Summary of the Duong's decline model behaviour, assumptions, conditions and parameters

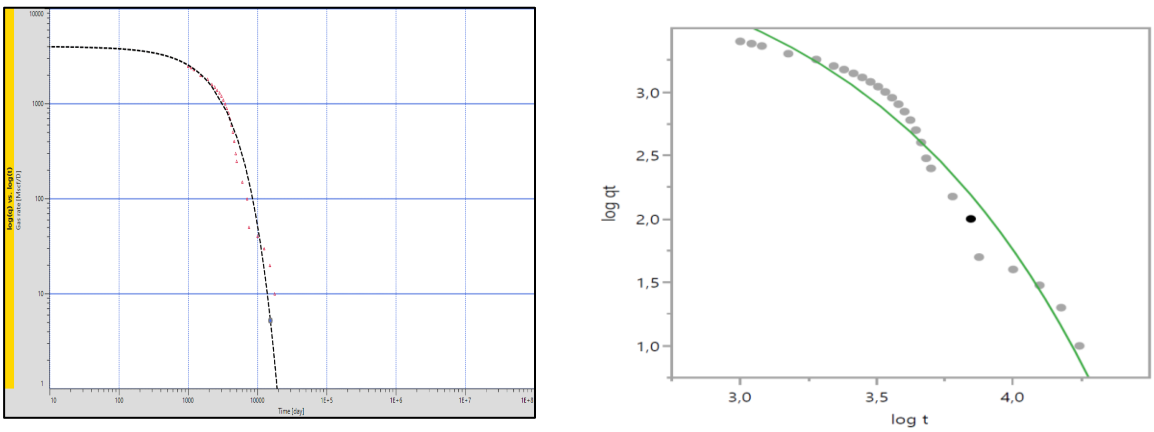
	Production Behaviour	**Assumptions**	**Conditions**	**Parameters**
Canon Shale Well	Transient Flow Regime (TFR	Very low permeability and long periods of transient flow The rate of decline approximated using a linear decline	1 < m < 2	a = 2.74 m = 1.53 q_i = 2492 EUR = 0.152 R^2 = 0.9402 RMSE = 0.0627
Marcellus Shale Well	Transient Flow Regime (TFR)	in the case of water breakthrough a and m increases	1 < m < 2	a = 3.03 m = 1.29 q_i = 2529 EUR = 16.14 R^2 = 0.8999 RMSE = 0.2261

6.2.3. Power Law Exponential Decline Model (PLE)

Ilk et al. (2008) presented the PLE, which is an extension of the exponential Arps formula for the decline degree in shale reservoirs. This model was developed precisely for SGR and approximates the rate of decline with a power law decline. The PLE model matches production data in both the transient and boundary-dominated regions without being hypersensitive to remaining reserve estimates (Kanfar et al., 2012). Seshadri and Mattar (2010) presented that the PLE model can model transient radial and linear flows, while Kanfar and Wattenbarger (2012) proved that the model is reliable for linear flow, bilinear flow followed by linear flow, and linear flow followed by BDF, or bilinear flow, followed by linear flow and finished with BDF flow. Vanorsdale (2013) deduced that when the flow regime changes throughout the initial 10 years of the well, the PLE model will yield a very optimistic recovery. The model characterizes the decline rate by infinite time, D_∞ which is defined as a "loss ratio" (which is assumed to be constant from Arps (1945). The production rate is derived as per Equations (2.6; 2.7 and 2.8).



(a)



(b)

Figure 6.3: KAPPA-Citrine and JMP graphs representing the PLE plot for (a) Canon Shale Well and (b) Marcellus Shale Well, using estimated production data

Based on **Figure 6.3**, the rate of decline is assumed to follow a power law decline. As in the other two cases, the production data points do not appear to fall on the fitting curve (Refer to Table A1 and A2 in Appendix A for data). A *log qt vs. log t* yielded R^2 values were 0.9724 (Canon Well) and 0.9586 (Marcellus Shale), respectively. The summary of the PLE model simulation can be seen in **Table 6.3**.

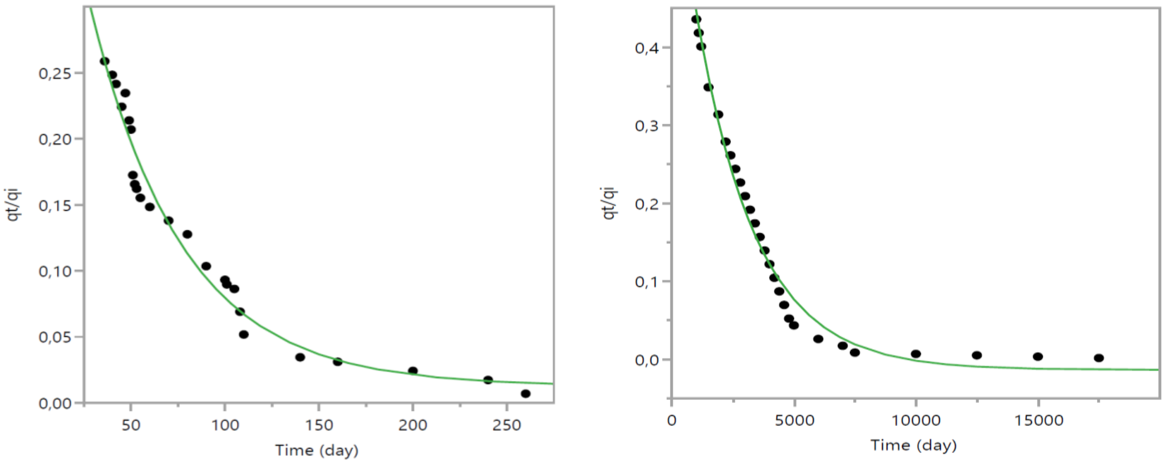
Table 6.3: Summary of the PLE decline model behaviour, assumptions, conditions and parameters

	Production Behaviour	Assumptions	Conditions	Parameters
Canon Shale Well	BDF and TFR	Approximates the rate of decline with a power law decline	D_i changes at early stages while D_∞ is constant at late time	n = 0.524 D_i = 0.0000681 q_i = 4018 EUR = 0.101 R^2 = 0.9724 RMSE = 0.0716
Marcellus Shale Well	BDF and TFR			n = 0.005 D_i = 0.0000431 q_i = 10812 EUR = 9.09 R^2 = 0.9586 RMSE = 0.1483

6.2.4. The Arps-Duong's-Power Law Hybrid Decline Model

The first proposed method incorporates the three DCA models, namely Arps, Duong's and PLE models. The Arps model only considers BDF while Duong's and PLE models consider TFR. The PLE model also considers BDF and has been specifically developed for SGRs. Hence, by combining the three models the limitations from each are presumed to be minimised or eliminated. Please refer to equation (5.3).

A plot of $\frac{qt}{qi}$ vs. t **Figure 6.4** provides the best fit for the model.



(a) **(b)**

Figure 6.4: A plot of $\frac{qt}{q_i}$ vs. t for (a) Canon Shale Well and (b) Marcellus Shale Well using estimated production data

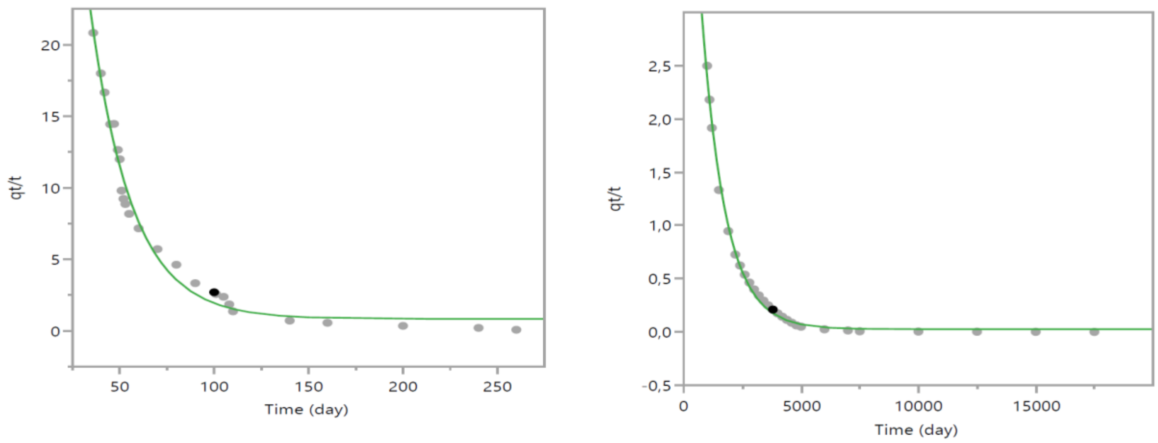
Figure 6.4 shows that the rate of decline initially follows a hyperbolic decline and then switches to an exponential decline. In this instance, it appears more of the data points fall on the fitting curve (Refer to Table A1 and A2 in Appendix A for data). The R^2 values obtained for the plot ($\frac{qt}{qi}$ vs.t) were 0.9686 (Canon Well) and 0.9864 (Marcellus Well), respectively. The summary of the Arps-Duong-Power Law hybrid decline model simulation can be seen in **Table 6.4.**

Table 6.4: Summary of the Arps-Duong-Power Law hybrid decline model behaviour, assumptions, conditions and parameters

	Production Behaviour	Assumptions	Conditions	Parameters
Canon Shale Well	BDF and TFR	Decline rate undergoes a switch point from a hyperbolic decline to an exponential decline	$0 \leq b \leq 1$ D_i changes at early stages and D_∞ becomes constant at late time	n = 0.755 D_i = 1.22 q_i = 1860 b = 0.188 EUR = 0.101 R^2 = 0.9686 RMSE = 0.4464
Marcellus Shale Well	BDF and TFR			n = 0.005 D_i = 0.0430 q_i = 7408 b = 0.00 EUR = 10.90 R^2 = 0.9864 RMSE = 0.0169

6.2.5. The Arps-Duong's Hybrid Decline Model

The second proposed model incorporates the two developed DCA models. Arps only considers BDF while Duong's considers TFR. Hence, both these flow regimes will be taken into account when combining these two models. The equation is given in (5.5). A plot of $\frac{qt}{t}$ vs. t **Figure 6.5** provides the best fit for the model.



(a) **(b)**

Figure 6.5: A plot of $\frac{qt}{t}$ vs. t for (a) Canon Shale Well and (b) Marcellus Shale Well using estimated production data

Based on **Figure 6.5**, the rate of decline follows a mechanistic growth decline. It appears that most of the data points fall on the fitting curve (Refer to Tables A1 and A2 in Appendix A for data). The R^2 values were 0.9846 (Canon Well) and 0.9958 (Marcellus Well) for the plot of $\frac{qt}{t}$ vs. t. The summary of the Arps-Duong hybrid decline model simulation can be seen in **Table 6.5**

Table 6.5: Summary of the Arps-Duong hybrid decline model behaviour, assumptions, conditions and parameters

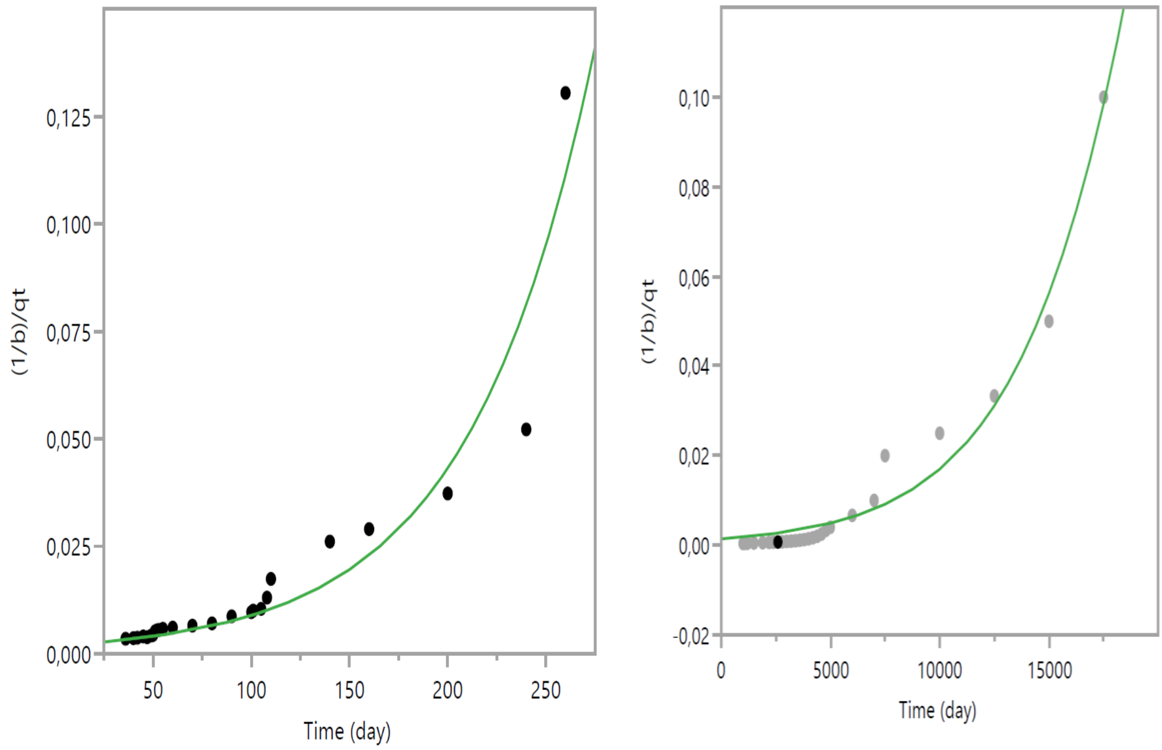
	Production Behaviour	Assumptions	Conditions	Parameters
Canon Shale Well	BDF and TFR	Approximates the rate of decline with a mechanistic growth decline	$0 \leq b \leq 1$	$D_i = 12.90$ $q_i = 6535$ b = 0.690 EUR = 0.161 $R^2 = 0.9846$ RMSE = 0.8154
Marcellus Shale Well	BDF and TFR		$0 \leq b \leq 1$	$D_i = 0.045$ $q_i = 2249$ b = 0.00 EUR = 12.28 $R^2 = 0.9958$

	RMSE = 0.3933

6.2.6. The Arps-Power Law Exponential Hybrid Decline Model

The third proposed model incorporates the Arps and PLE Models. These models consider BDF and TFR flows. Since the PLE model was developed specifically for SGRs, it would be advantageous to evaluate these two models combined, due to both being simple Equations to use. The equation is given in (5.6).

A plot of $\frac{\frac{1}{b}}{qt}$ vs. t **Figure 6.6** provides the best fit for the model.



(a) **(b)**

Figure 6.6: A plot of $\frac{\frac{1}{b}}{qt}$ vs. t for (a) Canon Shale Well and (b) Marcellus Shale Well using estimated production data

From **Figure 6.6**, the rate of decline follows a logistic growth decline. It appears the data points fall within the fitting curve. However, there are a few outliers observed for this model.

The plot from **Figure 6.6** yielded R^2 values were 0.9157 (Canon Well) and 0.9747 (Marcellus Shale). The summary of the Arps-Power Law hybrid decline model simulation can be seen in **Table 6.6**.

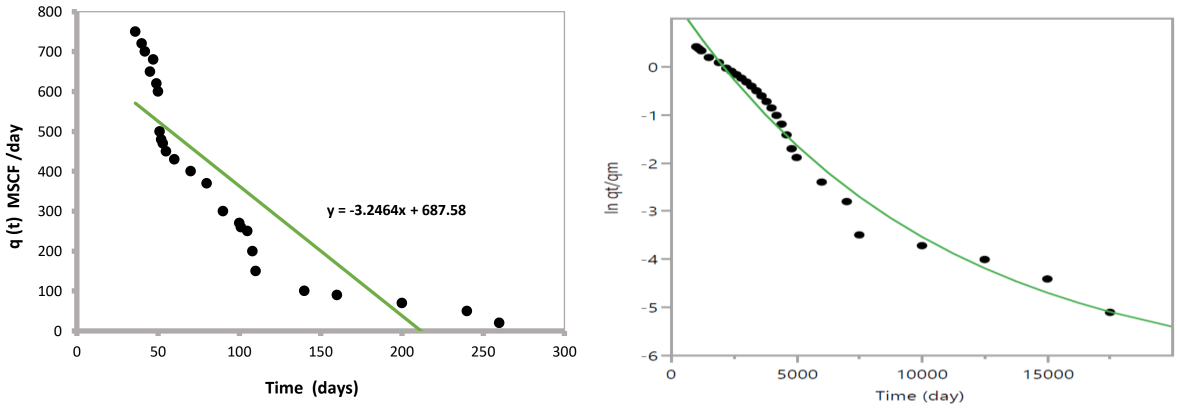
Table 6.6: Summary of the Arps-PLE hybrid decline model behaviour, assumptions, conditions and parameters

	Production Behaviour	Assumptions	Conditions	Parameters
Canon Shale Well	BDF and TFR	Approximates the rate of decline with a logistic growth decline	$0 \le b \le 1$ D_i changes at early stages and D_∞ becomes constant at late time	n = 1.00 D_i = 0.798 q_i = 1429 b = 0.000 EUR = 0.0904 R^2 = 0.9157 RMSE = 0.0081
Marcellus Shale Well	BDF and TFR			n = 0.005 D_i = 0.0139 q_i = 359 b = 0.0087 EUR = 1.44 R^2 = 0.9747 RMSE = 0.0036

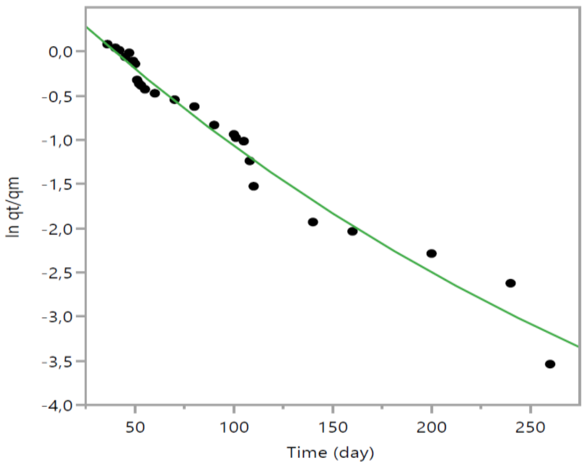
6.2.7. The Duong-Power Law Exponential Hybrid Decline Model

The fourth proposed model incorporates the Duong's and PLE Models. These models both consider TFR. The equation is in (5.7).

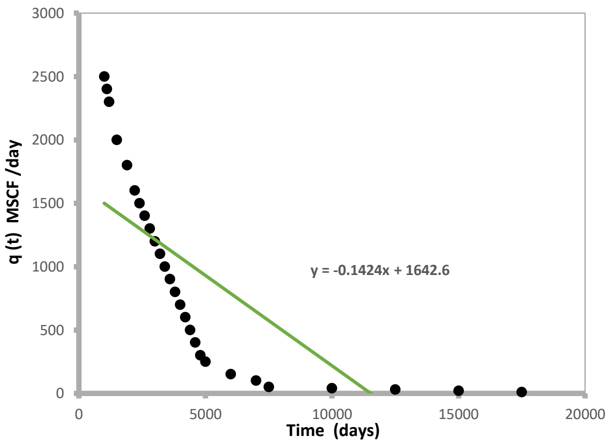
A plot of $\frac{In\ qt}{qm}$ vs. t **Figure 6.7** provides the best fit for the model.



(a) (b)



(c)



(d)

Figure 6.7: A plot of $\frac{In\,qt}{qm}$ vs. t for (a) Canon Shale Well and (b) Marcellus Shale Well using estimated production data with (c) and (d) being the excel plots used to determine qm

Based on **Figure 6.7**, the rate of decline follows a mechanistic growth decline. It appears that the data points do not all fall within the fitting curve. The R^2 values for the plot of $\frac{In\,qt}{qm}$ vs. t were 0.9718 and 0.9709, respectively. The summary of the Duong-Power Law hybrid decline model simulation can be seen in **Table 6.7**.

Table 6.7: Summary of the Duong's-PLE hybrid decline model behaviour, assumptions, conditions and parameters

	Production Behaviour	Assumptions	Conditions	Parameters
Canon Shale Well	BDF and TFR	Approximates the rate of decline with a mechanistic growth decline	$m > 1$ and D_∞ becomes constant at late time	$n = 0.781$ $D_i = 0.0000564$ $m = 1.59$ $q_i = 1020$ $qm = 688$ EUR = 0.106 $R^2 = 0.9718$ RMSE = 0.1669
Marcellus Shale Well	BDF and TFR			$n = 0.207$ $D_i = 0.0000430$ $m = 1.41$ $q_i = 29181$ $qm = 1643$ EUR = 20.84 $R^2 = 0.9709$ RMSE = 0.2865

6.2.8. Autoregressive Integrated Moving Average (ARIMA) Model

In order to evaluate the best fit for the ARIMA model a number of node scenarios were evaluated and the (3,1,3) and the (2,1,2) were selected to give the best forecast values because they were the lowest MSE and highest adjusted R^2 for the Canon Shale Well and Marcellus Shale Well, respectively. **Table 6.8** indicates the best results for the ARIMA model, which is highlighted in bold.

Table 6.8: Statistical results for the varying p,d,q for the ARIMA model obtained using JMP software

ARIMA	MSE (Canon Shale Well)	Adjusted R^2 (Canon Shale Well)	MSE (Marcellus Shale Well)	Adjusted R^2 (Marcellus Shale Well)
(0,0,0)	19.8	0.0000	66.7	0.0000
(1,1,1)	1.8	0.9880	2.4	0.9960
(1,2,1)	2.5	0.9780	2.7	0.9950
(1,3,1)	2.9	0.9670	3.8	0.9890
(2,1,1)	1.8	0.9880	2.5	0.9960
(2,1,2)	1.6	0.9910	**2.3**	**0.9960**
(3,1,3)	**1.2**	**0.9940**	3.6	0.9950

The graphical representation of the ARIMA models for the Canon Shale Well (3,1,3) and Marcellus Shale Well (2,1,2) can be seen in **Figure 6.8**.

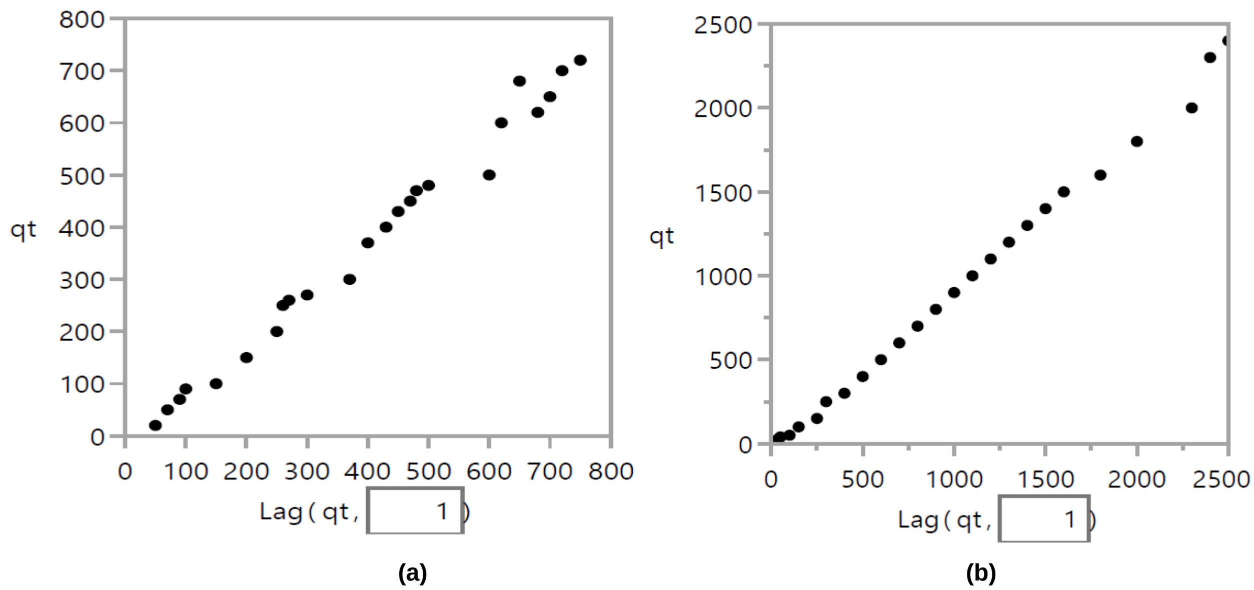


Figure 6.8: An ARIMA lag plot of qt vs. t for (a) Canon Shale Well and (b) Marcellus Shale Well

6.2.9. Artificial Neutral Network (ANN) Model

To choose the best algorithm for the model, a number of hidden nodes and layers were changed. In this study, the number of neurons was varied between 1 and 15. It has been highlighted in the literature that accuracy can be increased by increasing the number of nodes and layers (Khashei et al., 2009). The accuracy of the ANN-based estimation model was evaluated by various criteria, including mean percentage error (MPE), mean absolute percent error (MAPE) and correlation coefficient (R^2) between the actual and predicted values. The MPE, MAPE and R^2 are defined as follows:

$$MPE = \sum_{t=1}^{n} \frac{A_t - Pred_t}{A_t} \tag{6.14}$$

$$R^2 = \left(\frac{\sum_{t=1}^{n} (Pred_t - \overline{Pred_t})(A_t - \overline{A_t})}{\sum_{t=1}^{n} (Pred_t - \overline{Pred_t})^2 \sum_{t=1}^{n} (A_t - \overline{A_t})^2} \right) \tag{6.15}$$

where $Pred_t$ is the predicted value obtained from the neural network model and A_t is the actual value. $\bar{A}_t$ and $\overline{Pred}_t$ are the average of the actual and predicted values, respectively. **Table 6.9** summaries the performance output for the different number of neurons in the hidden layer for the Canon Shale Well and Marcellus Shale Well.

Table 6.9: Summary of the performance output for the different number of neurons in the hidden layer for the Canon Shale Well and Marcellus Shale Well

	Canon Shale Well		Marcellus Shale Well	
No. of Neurons	**MPE**	**R^2**	**MPE**	**R^2**
1	-3.85	0.9727	-3.33	0.9908
2	-3.25	0.9726	14.25	0.9954
3	-1.87	0.9730	12.48	0.9916
4	**-1.06**	**0.9743**	**0.44**	**0.9965**
5	-5.10	0.9855	21.14	0.9964
6	-10.26	0.9667	20.35	0.9999
7	-2.26	0.9864	5.04	0.9999
8	-2.85	0.9886	-3.72	1.0000
9	-8.67	0.9930	-5.42	1.0000
10	-4.37	0.9961	-1.71	1.0000
11	-7.26	0.9958	-17.96	0.9995
12	-3.49	0.9982	-5.11	1.000
13	-4.63	0.9972	-7.93	1.000
14	-4.06	0.9978	-6.39	1.000
15	-6.43	0.9912	2.89	1.000

In this study, for both the Canon Shale Well and Marcellus Shale Well a univariant input layer (a single input was used because of the data used in the study) and four nodes (1-4-1) highlighted in bold in **Table 6.9** (lowest MPE) gave the best model which is shown in **Figure 6.9**.

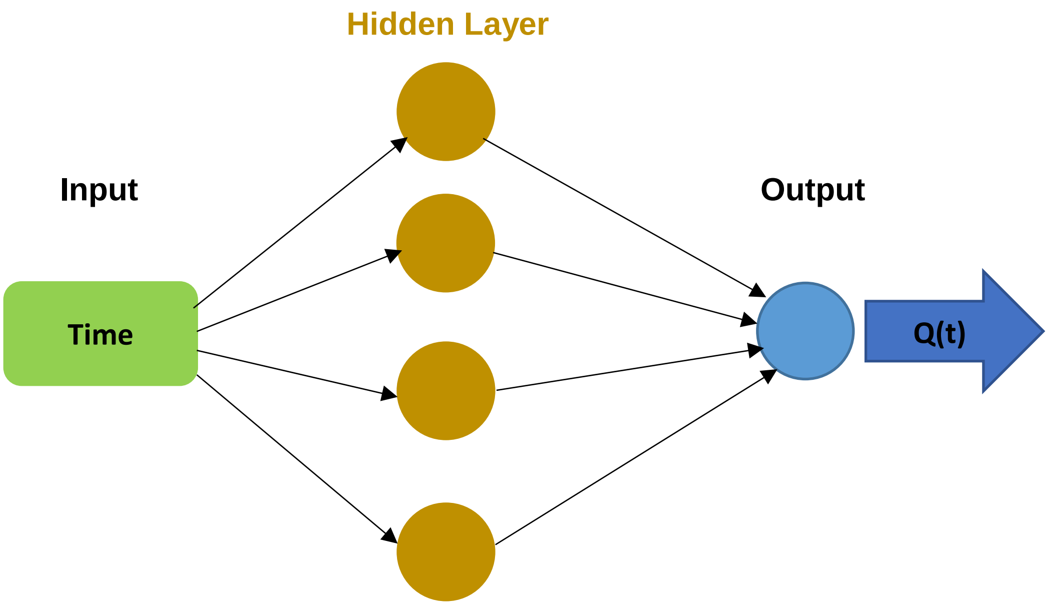


Figure 6.9: Univariant ANN model for both sets of production data applied in this study

6.2.10. ANN-ARIMA Hybrid Model

Zhang investigated the concept of the hybrid ANN-ARIMA model to obtain precise results as compared to using both models separately (Zhang, 2003). Numerous techniques, which explored the hybrid approach, have been used for many years to take advantage of the unique strengths of each type of models. The objective of merging the models is the notion that a single model cannot define all the specifics of time series (Tan et al., 2018).

N_t (Refer to equation 2.26) is obtained from the predicted values of the ANN model while $\hat{L}_t$ is the forecasted value from ARIMA based on the residual values. Regarding the hybrid model, the parameters used for the ARIMA, and ANN were used i.e., (2,1,2), (3,1,3) and a univariant ANN model, respectively. In the first step, ANN is used to predict the production rate, then the residuals e_t being produced are provided to ARIMA to predict the error. In the second step, the predicted production data by ANN is summed with the error produced by ARIMA to give the final predicted values. The error obtained from the ARIMA model was 22.1% (Canon Shale Well) and 39.7% (Marcellus Shale Well).

6.3. Conclusion

The section covered the simulation of the production data from the Canon and Marcellus Shale Well, using KAPPA-Citrine and JMP software. The simulation results indicated for the Arps, Duong and Power Law decline models showed that not all the production data points fall within the fitting curve, which was confirmed by the R^2 values. Conversely, for the hybrid models (order of rank based on the R^2 values), the Arps-Duong, Arps-Duong's-Power Law, Duong-Power Law and Arps-Power Law, the data appears to fall on the fitting curve better than the single models. The R^2 values obtained for the models showed higher values (Refer to **Table 6.10** for the summary of R^2 values for all models discussed).

In order to evaluate the best fit for the ARIMA model, a number of scenarios were evaluated and the ARIMA scenario (2,1,2) was selected to give the best forecast value for the Marcellus Well; due to having the lowest MSE of 4.82, a low BIC of 168.23 and highest adjusted R^2 of 0.9790. For the Canon, ARIMA scenario (3,1,3) was selected to give the best forecast value, due to having the lowest MSE of 6.31, a low BIC of 226.0 and the highest adjusted R^2 of 0.9937.

To choose the best algorithm for the ANN model, a number of hidden nodes and layers are changed. In this study, the number of neurons were varied between 1 to 15. The accuracy of the ANN-based estimation model was evaluated using the MPE and R^2 between the actual and predicted values. In this study, for both the Canon Shale Well and the Marcellus Shale Well, a univariant input layer and four nodes (1-4-1) gave the best model.

The ten models evaluated during the simulation process is summarised in **Table 6.10**. The next chapter will use the information obtained in the simulation process to evaluate the accuracy and validate the models.

Table 6.10: Summary of the data for each of the models during the simulation process

Model	Equation	Production Behaviour	Parameter Results	
			Canon Shale Well	**Marcellus Shale Well**
Arps Decline Model	$q_t = \frac{q_i}{(1 + bD_i t)^{\frac{1}{b}}}$	BDF	*b* = 0.383 D_i = 3.53 q_i = 2183 EUR = 0.100 R^2 = 0.9717 RMSE = 0.0725	*b* = 0.00 D_i = 0.042 q_i = 3864 EUR = 9.24 R^2 = 0.9692 RMSE = 0.1279
Duong's Decline Model	$q(t) = q_i t(a, m) + q_\infty$	TFR	*a* = 2.74 *m* = 1.53 q_i = 2492 EUR = 0.152 R^2 = 0.9402 RMSE = 0.0627	*a* = 3.03 *m* = 1.29 q_i = 2529 EUR = 16.14 R^2 = 0.8999 RMSE = 0.2261
PLE Decline Model	$q(t) = \hat{q}_i e^{[-D_\infty t - \hat{D}_i t^{\hat{n}}]}$	BDF and TFR	*n* = 0.524 D_i = 0.0000681 q_i = 4018 EUR = 0.101 R^2 = 0.9724 RMSE = 0.0716	*n* = 0.005 D_i = 0.0000431 q_i = 10812 EUR = 9.09 R^2 = 0.9586 RMSE = 0.1483
Arps-Duong's-PLE Hybrid Model	$\frac{qt}{qi} = t\,(-D_\infty - D_i \hat{n}) - In\,\frac{b+1}{b}$	BDF and TFR	*n* = 0.755 D_i = 1.22 q_i = 1860 *b* = 0.188 EUR = 0.101 R^2 = 0.9686 RMSE = 0.4464	*n* = 0.005 D_i = 0.0430 q_i = 7408 *b* = 0.00 EUR = 10.90 R^2 = 0.9864 RMSE = 0.0169

Arps-Duong Hybrid Model	$qt = [\frac{qt}{t}][1 + bD_i]^{-\frac{1}{b}}$	BDF and TFR	D_i = 12.90 q_i = 6535 b = 0.690 EUR = 0.161 R^2 = 0.9846 RMSE = 0.8154	D_i = 0.045 q_i = 2249 b = 0.00 EUR = 12.28 R^2 = 0.9958 RMSE = 0.3933
Arps-PLE Hybrid Model	$t[-D_\infty - D_i\hat{n}] = \frac{-\frac{1}{b}}{qt} In(1 + bD_i)$	BDF and TFR	n = 1.00 D_i = 0.798 q_i = 1429 b = 0.000 EUR = 0.0904 R^2 = 0.9157 RMSE = 0.0081	n = 0.005 D_i = 0.0139 q_i = 359 b = 0.0087 EUR = 1.44 R^2 = 0.9747 RMSE = 0.0036
Duong-PLE Hybrid Models	$\frac{In\ qt}{qm} = t\ [-D_\infty - D_i\hat{n}]$	BDF and TFR	n = 0.781 D_i = 0.0000564 m = 1.59 q_i = 1020 qm = 688 EUR = 0.106 R^2 = 0.9718 RMSE = 0.1669	n = 0.207 D_i = 0.0000430 m = 1.41 q_i = 29181 qm = 1643 EUR = 20.84 R^2 = 0.9709 RMSE = 0.2865
ARIMA Model	$Y_t = \propto_1 Y_{t-1} + \varepsilon_t$ $Y_t = Y_{t-1} + \varepsilon_t$ $Y_t = \varepsilon_t - \theta_1\varepsilon_{t-1}$	Linear	(2,1,2) RMSE = 0.00304	(3,1,3) RMSE = 0.0472
ANN Model	$Y_t = \propto_0 + \sum_{j-1}^{q} \propto_j g\left(\beta_0 j + \sum_{I=1}^{p} \beta_i j Y_t - i\right) + \varepsilon_t$	Non-linear	(1-4-1) RMSE = 0.0148	(1-4-1) RMSE = 0.0165
ANN-ARIMA Hybrid Model	$\hat{y}_t = \hat{L}_t + \hat{N}_t.$	Linear and Non-linear	(1-4-1) (2,1,2) RMSE = 0.2384	(1-4-1) (3,1,3) RMSE = 0.2049

Chapter 7

The Forecasting, Accuracy and Validation of Models

7.1. Introduction

This chapter will discuss the accuracy and validate each of the models, based on the data obtained during simulation. In order to validate the result of each forecasting model, which was discussed in the previous section, model validation is required, based on the different error analysis method.

The MSE; RMSE and R^2 between the actual and predicted data were calculated using JMP software for each model. The aim of the model validation is to find the best model which gives the least error and the best fit of the data. The best model is then selected to predict future shale gas demand.

7.2. Production Forecasting[13]

7.2.1. The Arps Decline Model

KAPPA-Citrine software was initially used for determining the parameters for the Arps decline model. The *b* values were found to be 0.383 and 0.00, while the D_i values were 3.53 and 0.042, respectively, for Cannon Well and Marcellus Shale Wells.

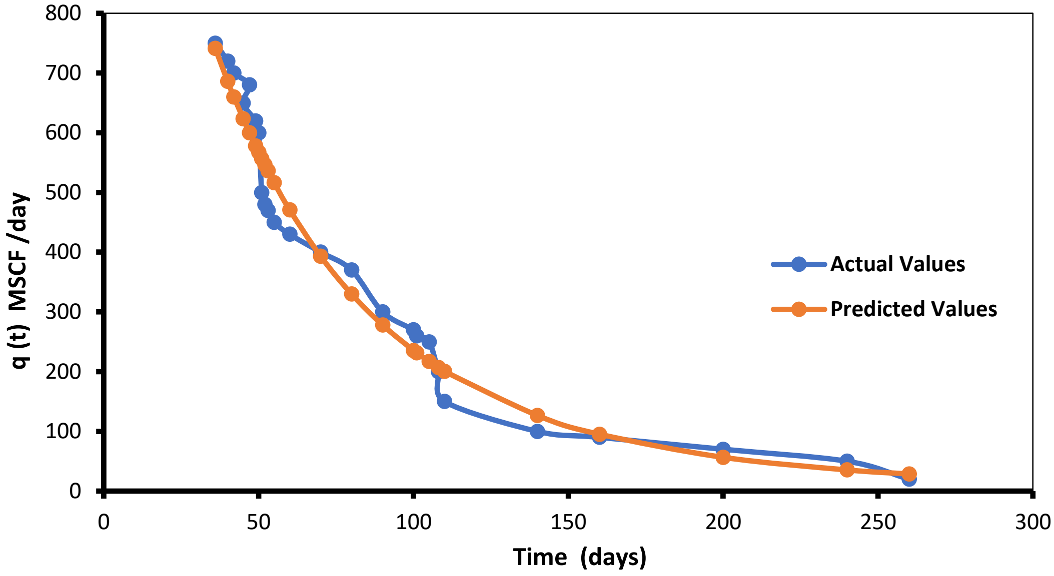
Subsequently, JMP software was used to construct the prediction model. The second step was to graph a semi log plot (*log q vs. t*) to determine the model forecasting Equation (7.1) and parameters.

The forecasting equation is given as follows:

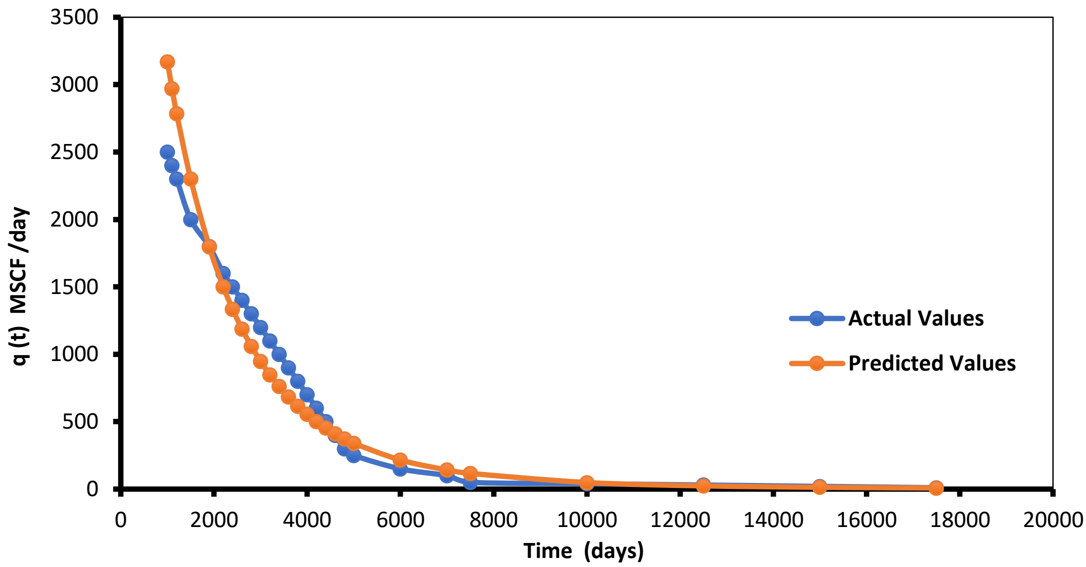
$$y = \frac{c}{1 + e^{(-ax^2 - b)}} \quad \textbf{(7.1)}$$

where *c* is the asymptote, *a* the growth rate, while *b* is the inflection point and *x* is time. The actual and forecasted flow rate values are shown graphically in **Figure 7.1**.

[13] Raw data can be found in Appendix A – Section A.3. Actual vs. Predicted Data



(a)



(b)

Figure 7.1: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Arps decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

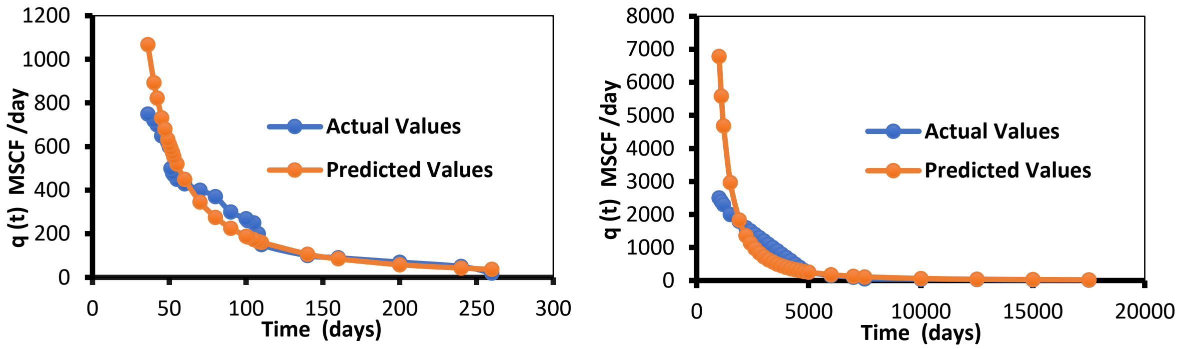
The results for the model appear in some instances to over-and in other instances to under-estimate the data. The results concur with the literature, which suggests that the weakness of the Arps decline model is overestimation of results. Tan et al. (2018) highlighted that although the Arps decline model is simple and fast, it often fails to accurately fit the decline curve of unconventional reservoirs. They further explained that the model often tends to overestimate the EUR for shale gas wells because it assumes that a BDF regime is evident. Paryani et al. (2018) concurred with these findings, explaining that the drainage area is not constant because the pressure pulse continues to spread from the fracture to other areas of the reservoir volume.

7.2.2. The Duong's Decline Model

The parameters for the Duong decline model were q_i = 2492, *a* = 2.74 and *m* = 1.53 for Canon Shale Well and were q_i = 2529, *a* = 3.03 and *m* = 1.29 for the Marcellus Shale Well. In this instance, a log–log linear plot (log *q* vs. log *t*) was used. The forecasting equation is given as:

$$y = bx + c \tag{7.2}$$

where *b* is the slope, *x* is time and *c* is the intercept. The actual and forecasted flow rate values can been seen in **Figure 7.2.**



(a) **(b)**

Figure 7.2: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production, using the Duong's decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

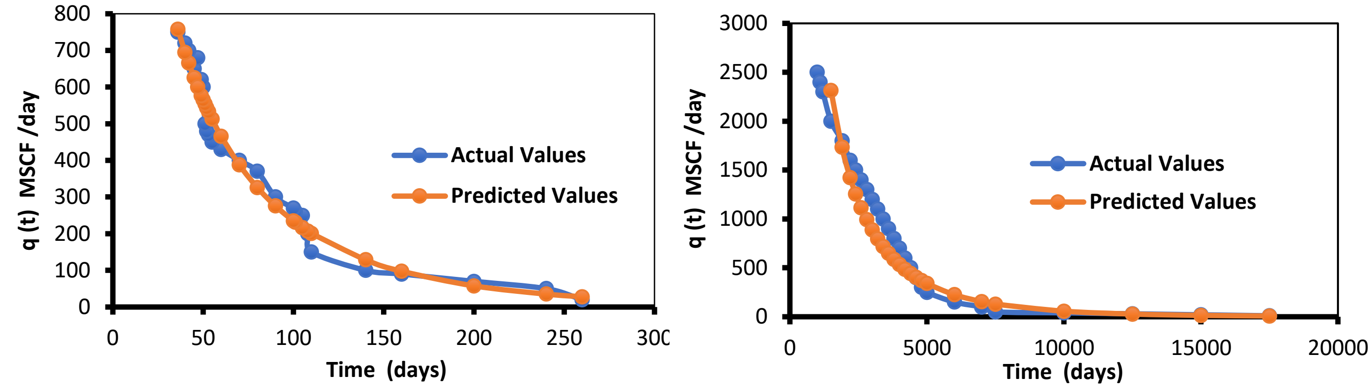
The results for the Duong decline model indicate an under and over-estimate of the data. Meyet et al. (2013) mentioned in their work that the Duong decline model tends to provide the most conservative results. This could also be attributed to the fact that the Duong decline model tends to be more accurate for linear flows and bilinear–linear flows (Kanfar et al., 2012). Paryani et al. (2018) found that the well fitted with 51% of the historical production data, and that the Duong decline model fits better with longer and less noisy historical production data.

7.2.3. The PLE Decline Model

The parameters used in the model for *n* and D_i are 0.524 and 0.0000681 for Canon Shale Well and *n* and D_i used were 0.005 and 0.0000431 for the Marcellus Shale Well. A log-log plot (*log q vs. log t*) was used in the model forecasting. The forecasting equation is given as:

$$y = a + be^{cx} \tag{7.3}$$

where *a* is the asymptote, *b* is the scale, *c* is the growth rate and *x* is time. The actual and forecasted values can be seen in **Figure 7.3**.



(a) **(b)**

Figure 7.3: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production, using the PLE decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

The results for the PLE decline model appear to underestimate the data overall, although the model considers BDF and TFR, which is an advantage of the model. Furthermore, the model was specifically developed for SGRs, hence it was assumed that the results would be better. This is comparative to the findings by Paryani et al. (2018); as based on their results the PLE consistently gave the lowest forecasts for all the models. It is therefore the most conservative method for production forecasting and reserves estimation (Paryani et al., 2018). Seshadri and Mattar (2010) concluded that for tight gas wells, the PLE decline model is complex and non-intuitive. The power law model can result in a non-unique solution due to four degrees of freedom, resulting from the four unknown parameters (Ali and Sheng, 2015).

7.2.4. The Arps-Duong-PLE Hybrid Decline Model

A plot of $\frac{qt}{qi}$ vs. *t* was used in the model forecasting. The parameter q_i used was 1860 and 7408 for the Canon Shale Well and Marcellus Shale Well, respectively. The forecasting equation is given as:

$$y = a + be^{cx} \quad \textbf{(7.4)}$$

where *a* is the asymptote, *b* is the scale, *c* is the growth rate and *x* is time. The actual and forecasted rates are graphically represented in **Figure 7.4**.

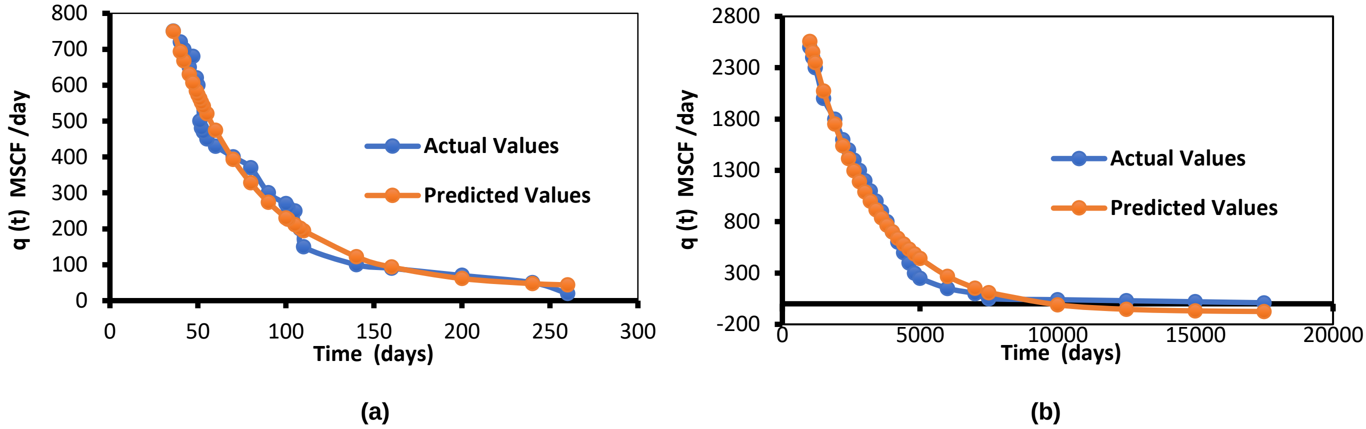


Figure 7.4: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production, using the Arps-Duong-PLE hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

Based on the results, the model appears to over-and underestimate the data. However, the gap between the actual and predicted results is minimised. This could be attributed to both BDF and TFR being considered. In addition, the conservative approach of Duong's and the PLE models, along with the inaccurate fitting of the Arps decline curve of unconventional reservoirs could be a contributing factor.

7.2.5. The Arps-Duong Hybrid Decline Model

A plot of $\frac{qt}{t}$ vs. *t* was used in the model forecasting. The forecasting equation is given as:

$$y = a(1 - e^{-cx}) \quad \textbf{(7.5)}$$

where *a* is the asymptote, *b* is the scale, *c* is the growth rate and *x* is time. The actual and forecasted values can be seen in **Figure 7.5**.

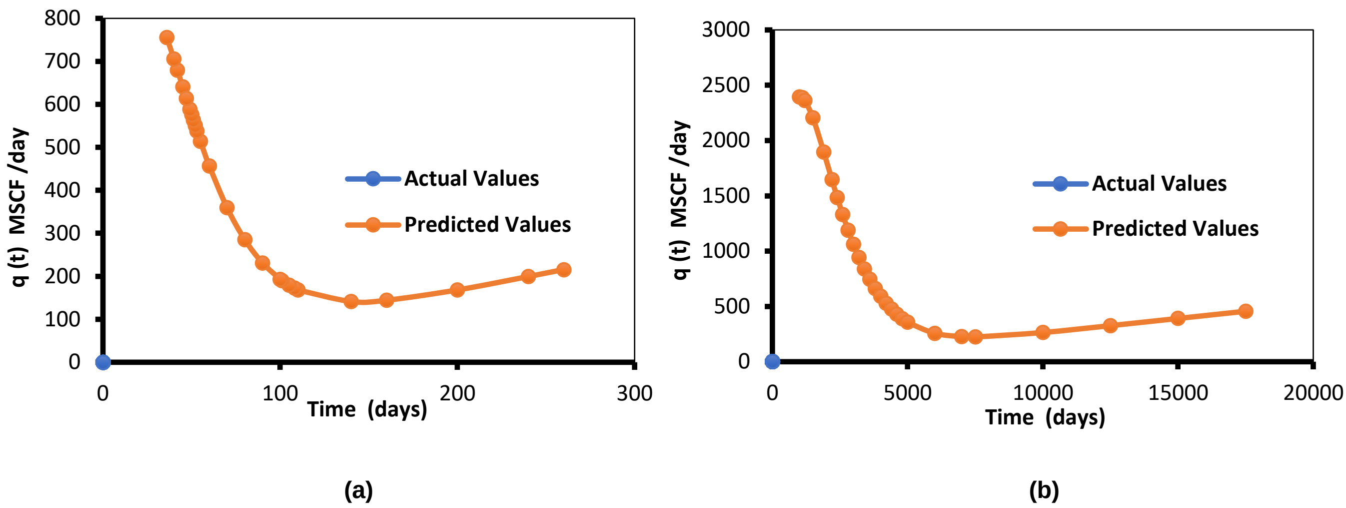


Figure 7.5: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Arps-Duong hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

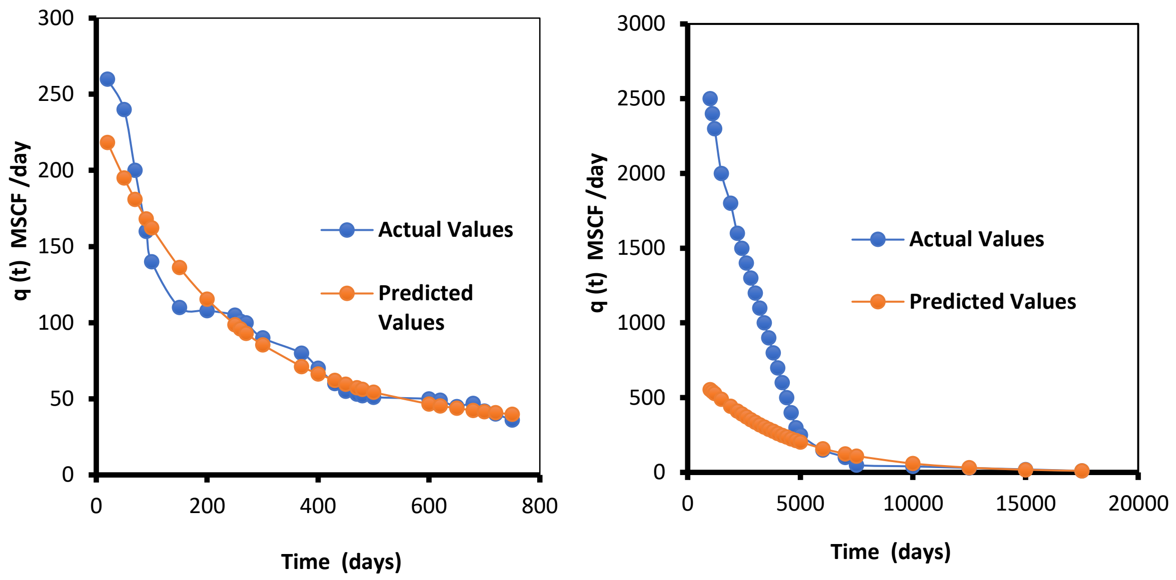
The predicted results for the model appear to be severely overestimated from the actual results in the latter stage of production. This would be the result of combining the drawbacks of the two models, which causes the elevated results observed. In line with this; firstly, most shale gas wells rarely reach the boundary-dominated flow regime, hence the Arps decline model cannot be applied directly to SGRs without significant modifications (Tan et al., 2018). Secondly, in the findings of Paryani et al. (2018), extremely high reserves estimates were occasionally observed with the Duong decline model. The results of Hu et al. (2018) concurred with these results, for the Austin Chalk wells, whereby the Duong decline model gave the highest weighted residual of production rate.

7.2.5. The Arps-PLE Hybrid Decline Model

A plot of $\frac{\frac{1}{b}}{qt}$ vs. t was used in the model forecasting. The forecasting equation is given as:

$$y = \frac{c}{1 + e^{(-ax-b)}} \qquad \textbf{(7.6)}$$

where *c* is the asymptote, *b* is the inflection point, *x* is time and *a* is the growth rate. The actual and forecasted values can be seen in **Figure 7.6**.



(a) **(b)**

Figure 7.6: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Arps-PLE hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

The results from the model initially appear to over-and underestimate the data prediction; however, the results tend to move closer to the actual values over time. This would be attributed to the reliability in the Arps decline model and the fact that the PLE decline model was developed precisely for SGR. Moreover, both flow regimes are considered and since most shale gas wells rarely reach the boundary-dominated flow regime, the results appear to move closer to the actuals when reaching the TFR. Hence, by combining the models, the overestimation of the predicted results is minimised over time.

7.2.5. The Duong-PLE Hybrid Decline Model

A plot of $\frac{In\, qt}{qm}$ vs. *t* was used in the model forecasting. The forecasting equation is given as:

$$y = a(1 - be^{-cx}) \qquad \textbf{(7.7)}$$

where *a* is the asymptote, *b* is the scale, *x* is time and *c* is the growth rate. The actual and forecasted values can be seen in **Figure 7.7**.

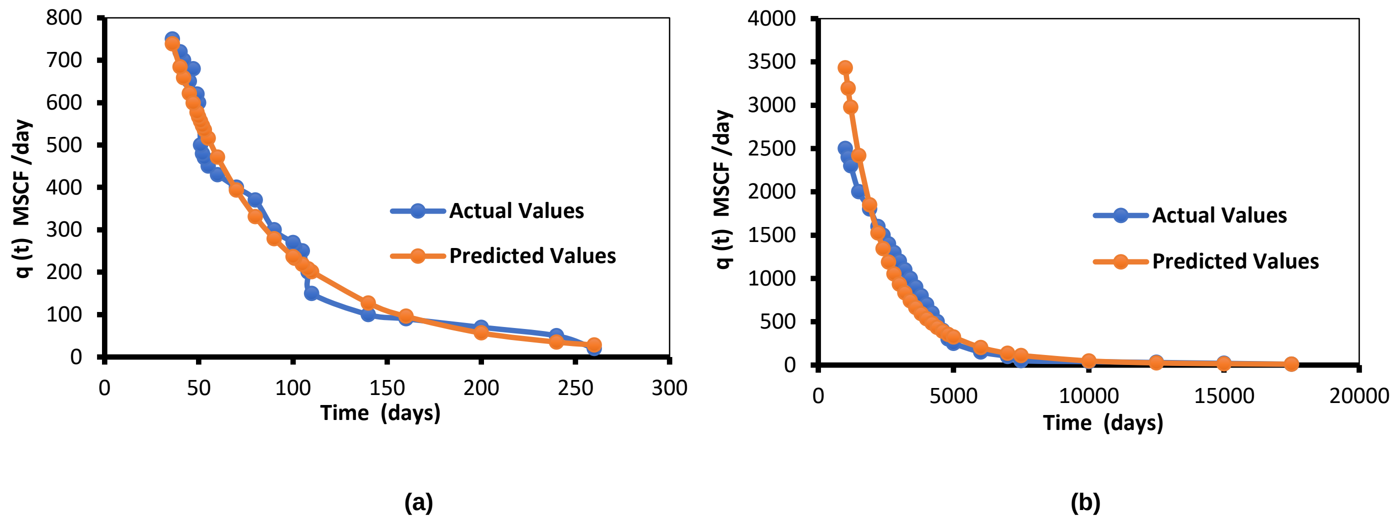


Figure 7.7: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Duong's-PLE hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

The trend of the results indicate an over-and underestimation. As mentioned by Vanorsdale (2013), the PLE and Duong's decline models will yield an optimistic recovery when the flow regime changes. This trend is evident in the results when combining the models.

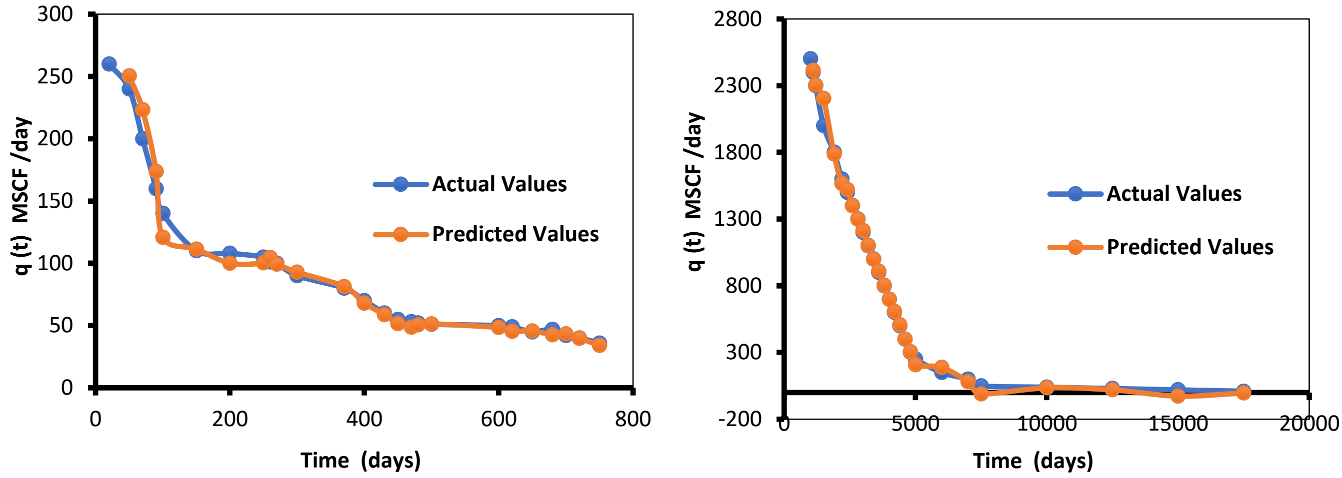
7.2.6. The ARIMA Model

As mentioned earlier under the simulation section, the best fit for the ARIMA model was a (3,1,3) for the Canon Shale Well and (2,1,2) for the Marcellus Shale Well. These parameters gave the best forecast values, since having the lowest MSE and highest adjusted R^2.

The best model is shown as follows:

$$Y_t = \theta_2 Y_{t-2} + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \varepsilon_t \qquad \textbf{(7.8)}$$

The actual and forecasted values can be seen in **Figure 7.8**.



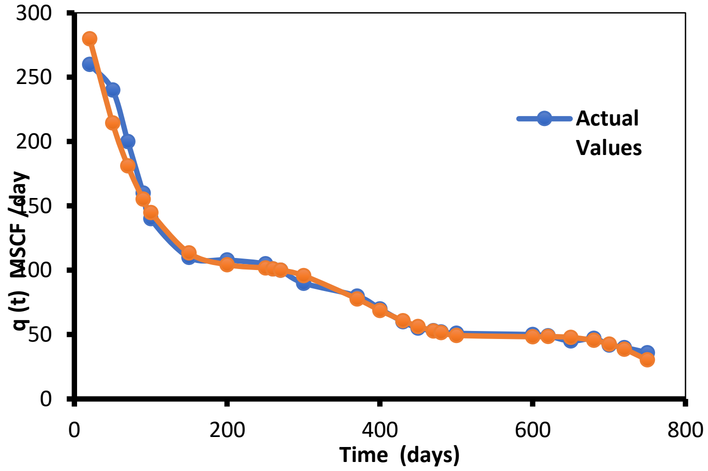
(a) (b)

Figure 7.8: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production, using the ARIMA model for (a) Canon Shale Well and (b) Marcellus Shale Well

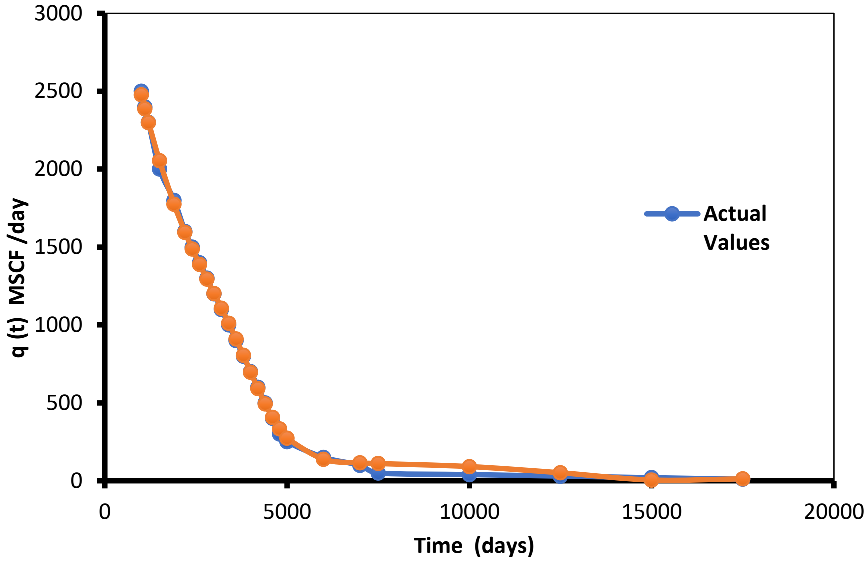
The predicted results from the model appear to follow a close trend to the actual values. Raymond (2007) suggested that ARIMA models have proved to be excellent short-term forecasting models for a wide variety of time series because short-term factors are expected to change slowly. This can explain the reason why the ARIMA fared well compared to the other models discussed so far.

7.2.7. The ANN Model

In the case of this study, a univariate input layer and four nodes gave the best model fit i.e. (1-4-1) for the production flow rate, over a period of time. The actual and forecasted values are graphically represented in **Figure 7.9.**



(a)



(b)

Figure 7.9: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the ANN model for (a) Canon Shale Well and (b) Marcellus Shale Well

The predicted results from the model appear to follow a very close trend to the actual values. Zhang (2003) indicated that neural networks are useful for modelling and predicting the properties of time series data.

Cybenko (1989) described neural networks as having a universal non-linear function and a relatively good degree of forecasting accuracy. In addition, according to Hill et al. (1996), neural network forecasting provides better results than traditional forecasting methods over monthly as well as quarterly periods.

7.2.8. ANN-ARIMA Hybrid Model

The steps employed by Ayub and Jafri (2020) was used to construct the ARIMA-ANN hybrid model. This involved a two-step process as follows:

In the first step, ANN is used to predict q_t and residuals e_t is produced and provided to ARIMA to predict the error. In the second step, the predicted q_t by ANN is summed with the error produced by the ARIMA model, to give the final predicted values. The equation (2.26) is given in Chapter 2.

The actual and forecasted values can be seen in **Figure 7.10**.

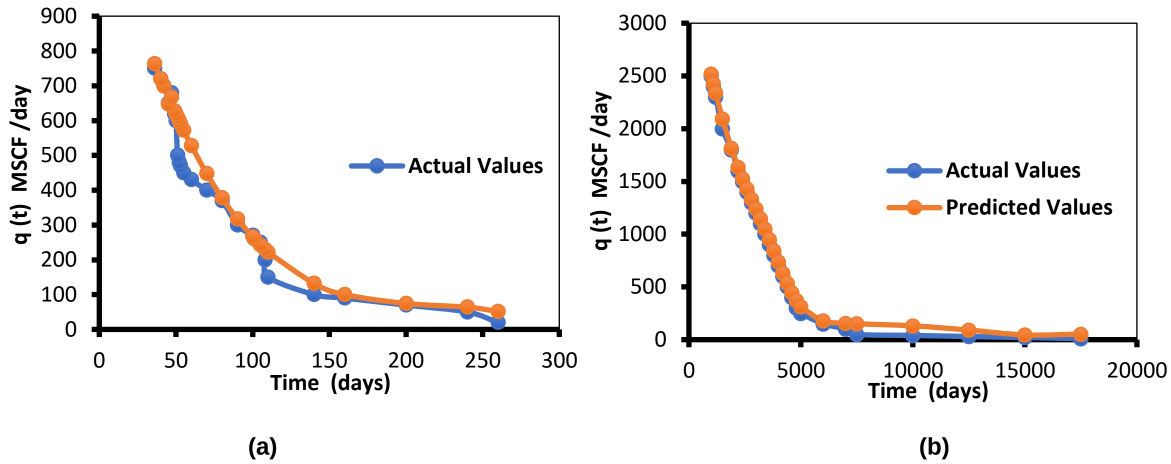


Figure 7.10: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the ANN-ARIMA hybrid model for (a) Canon Shale Well and (b) Marcellus Shale Well

The predicted results from the model appear to be overestimated compared to the actual values. The primary reason for the overestimation can be attributed to the error value obtained from ARIMA model. The higher the error value, the higher would be the predicted value. This finding appears to contradict what has been indicated in the literature. According to Faruk (2010), hybrid methods have a higher degree of accuracy than neural networks. Cybenko (1989) indicated in his work that hybrid models combine the advantages of ARIMA with respect to linear modelling and neural networks in terms of non-linear edge modelling (Cybenko, 1989). However, Taskaya-Temizel and Ahmad (2005) made reference in their work that in some circumstances, the single model approach can outperform hybrid models. This was observed during this study.

7.3. The Evaluation of the Hybrid Models

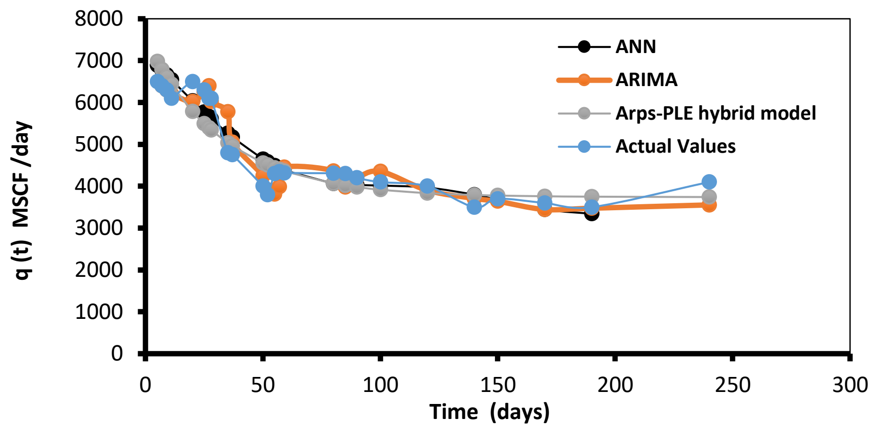
This study investigated the hybrid model approach to determine if this approach provides a higher accuracy than the single model. **Table 7.1** summarises the evaluation of the hybrid models. From the findings, it was observed that with the hybrid models there is over, under or both estimations of the predicted data vs. the actuals. Nevertheless, with the Arps-Duong-PLE hybrid decline model and the Arps-PLE hybrid decline model, the gap between the actual and predicted values is reduced. It appears from the evaluation the Arps-PLE hybrid decline model from all the models provides a moderate to low degree of over and under-estimation of the data to the actuals. The findings will be validated later in this chapter.

Table 7.1: Summary of the hybrid decline models

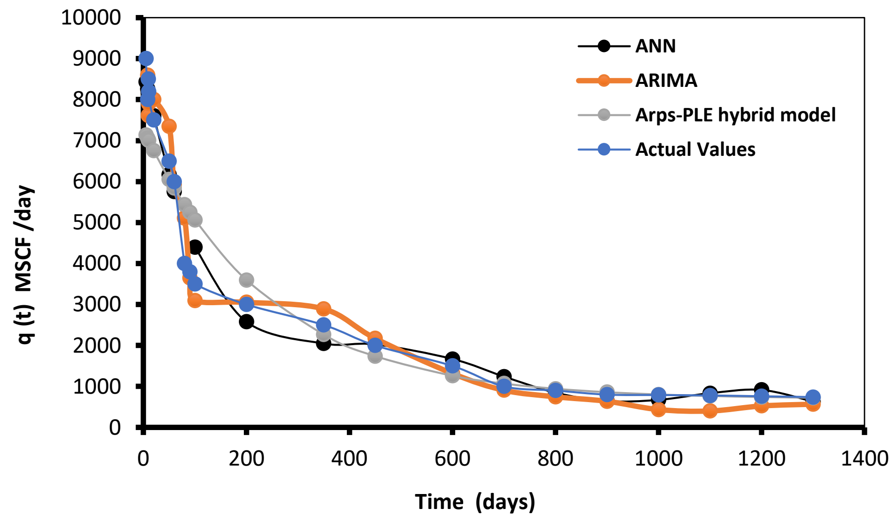
Model	Forecasting Equation	Forecasting Evaluation	Degree of over and underestimation
Arps-Duong-PLE hybrid decline model	$y = a + be^{cx}$	Resultant is over-and underestimated, the difference between the actual and predicted is reduced	High (over and under)
Arps-Duong hybrid decline model	$y = a(1 - e^{-cx})$	Resultant is overestimated	High (over)
Arps-PLE hybrid decline model	$y = \frac{c}{1 + e(^{-ax-b)}}$	Resultant is over- and underestimated however results move closer to the actuals	moderate (over), low (under)
Duong-PLE hybrid decline model	$y = a(1 - be^{-cx})$	Resultant is over- and underestimated	High to moderate (over), low (under)
ANN-ARIMA hybrid decline model	$\hat{y}_t = \hat{L}_t + \hat{N}_t$	Data is overestimated	Moderate to high (over)

7.4. The Evaluation of Model Accuracy

In order to assess the accuracy of the models, two sets of different production data were used to perform the evaluation. The estimated data was extracted from the work of Brantson et al. (2019) (Source Well #A in Chang-2 segment of Eastern Sichuan) and Tan et al. (2018) (sourced the data from the Barnett Shale Well gas well in Fort Worth Basin in Northeast Texas). Based on the forecasting evaluation, it was observed that the Arps-PLE hybrid decline model, ARIMA and ANN models provided the best forecasting accuracy. **Figure 7.11** illustrates the actual data vs. the predicted data for the ARIMA, ANN and Arps-PLE hybrid decline models.



(a)



(b)

Figure 7.11: Estimated production data to determine accuracy for (a) ARIMA vs. ANN vs. Arps-PLE hybrid decline model for well #A and (b) ARIMA vs. ANN vs. Arps-PLE hybrid decline model for Barnett Shale Well

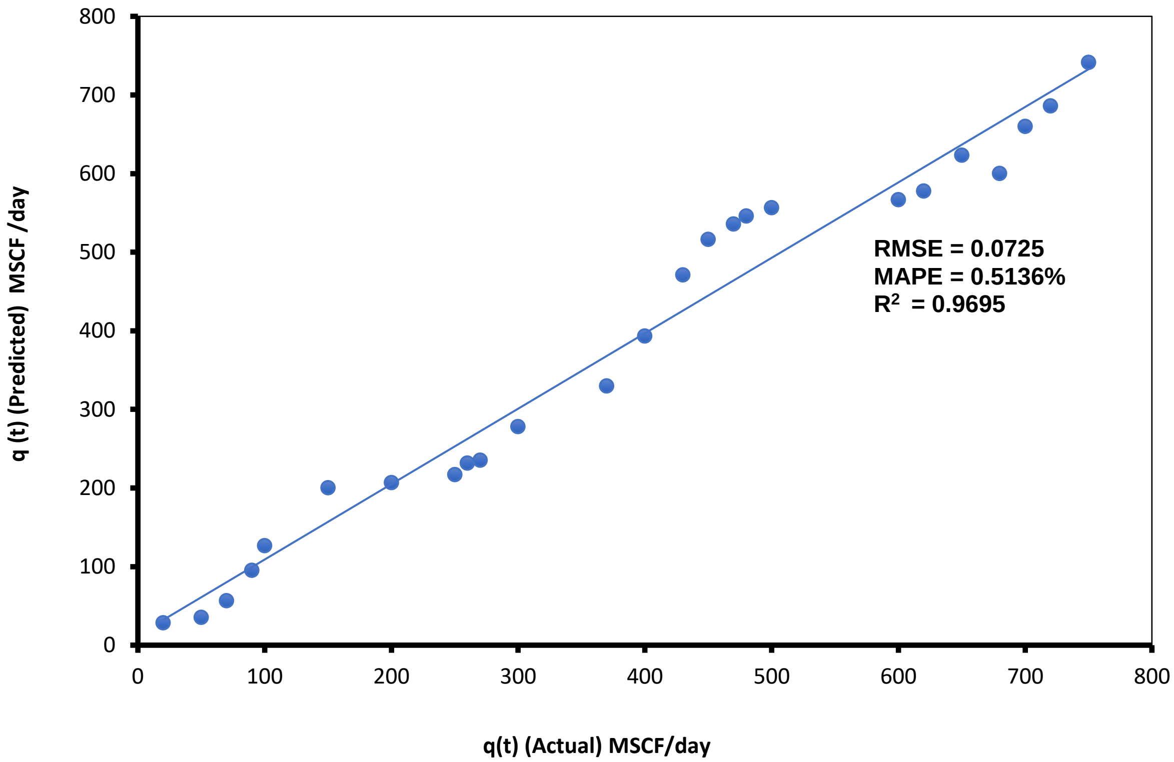
The results from graphs *a* and *b* indicate the consistence in the three models i.e., ARIMA, ANN and Arps-PLE hybrid decline models to predict the production data for the Well #A and Barnett Shale Wells, respectively. The predicted values in both cases were close to the actual values, as indicated in **Figure 7.11**.

7.5. The Validation of Models

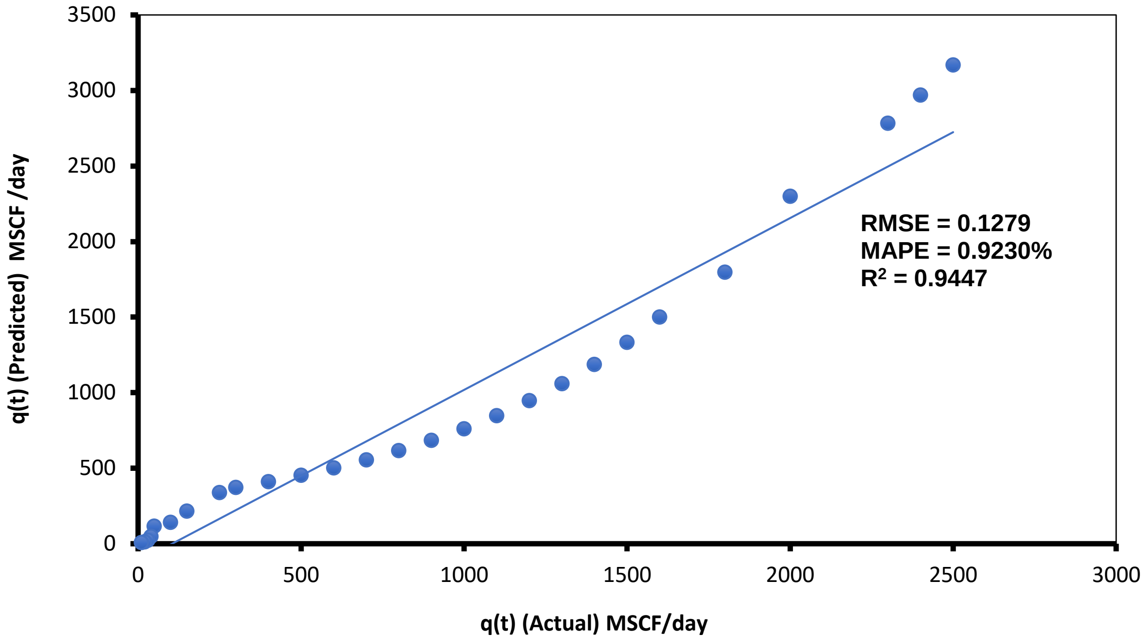
In order to determine the accuracy and validate the results obtained for the forecasting models, the RMSE, MAPE and R^2 between the actual and predicted were calculated, the formulas of which have been presented in Chapter 6. The data used for the model validation were from Canon Shale Well and Marcellus Shale Well, respectively.

7.5.1. The Arps Decline Model

Figure 7.12 illustrates the scatter plot of the predicted *q(t)* values by the Arps decline model compared to the experimental *q(t)* values. The RMSE, MAPE and R^2 values are illustrated in **Figure 7.12**. As can be seen from the figure R^2 values of 0.9695 and 0.9447 were obtained for the data sets, while the MAPE of 0.5136 and 0.9230% were calculated from the data.



(a)

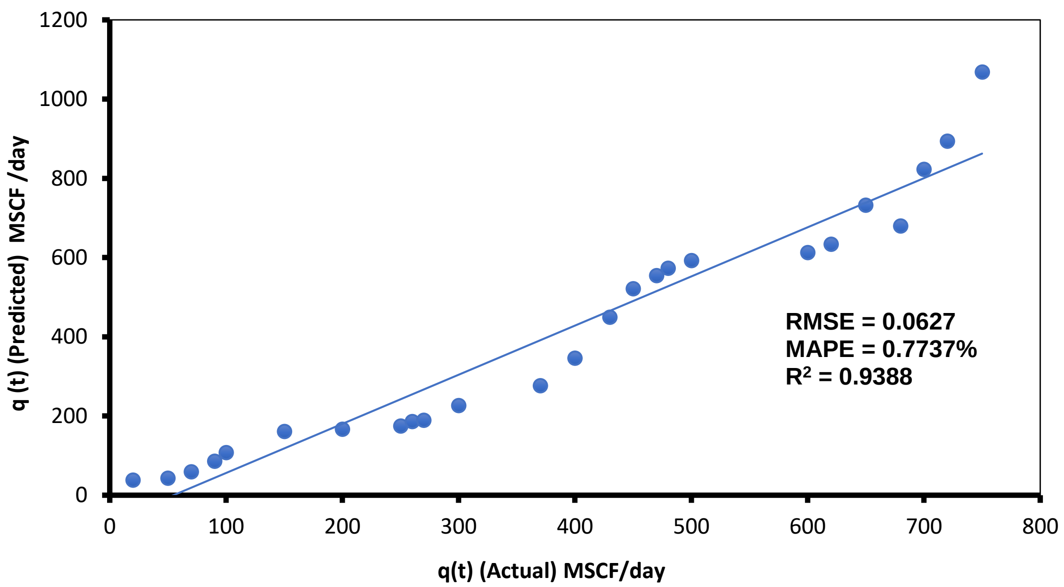


(b)

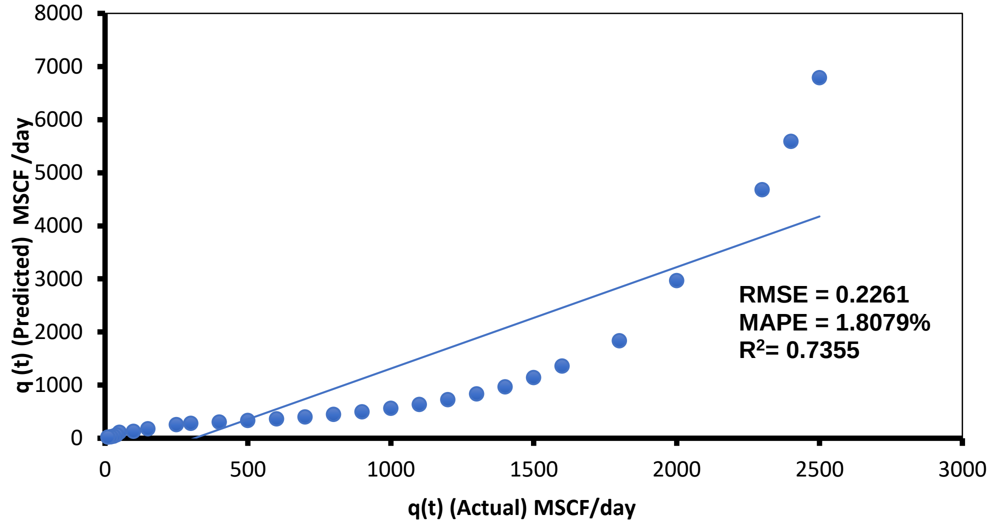
Figure 7.12: Predicted q(t) vs. actual q(t) for the Arps decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.2. The Duong's Decline Model

Figure 7.13 illustrates the scatter plot of the predicted *q(t)* values by the Duong's decline model against the experimental *q(t)* values. The figure shows R^2 values of 0.9388 and 0.7355 that were obtained for the data sets while the MAPE of 0.7737 and 1.8079% was calculated from the data.



(a)

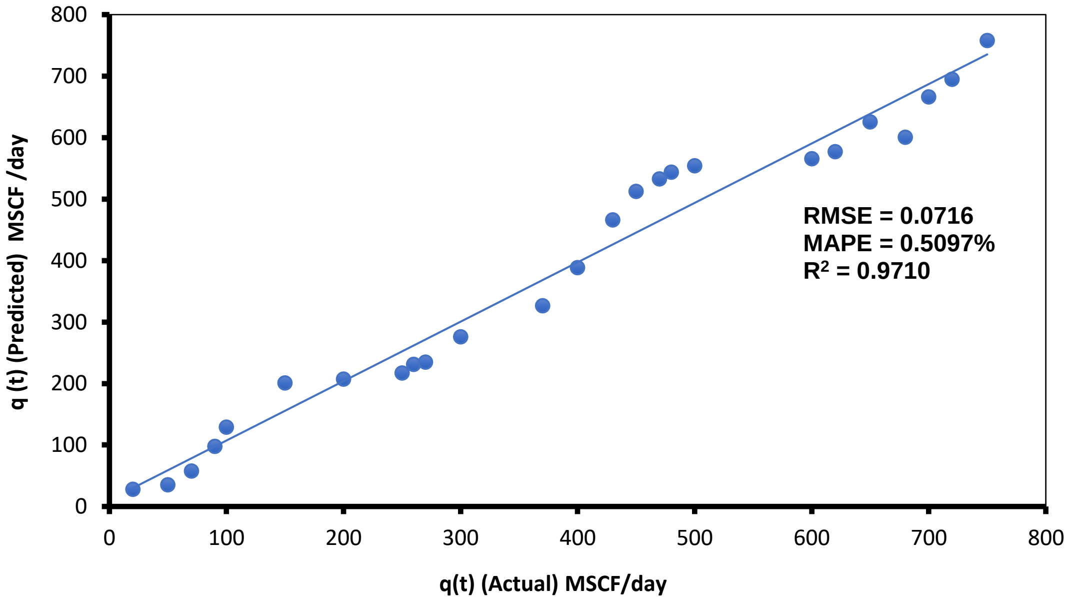


(b)

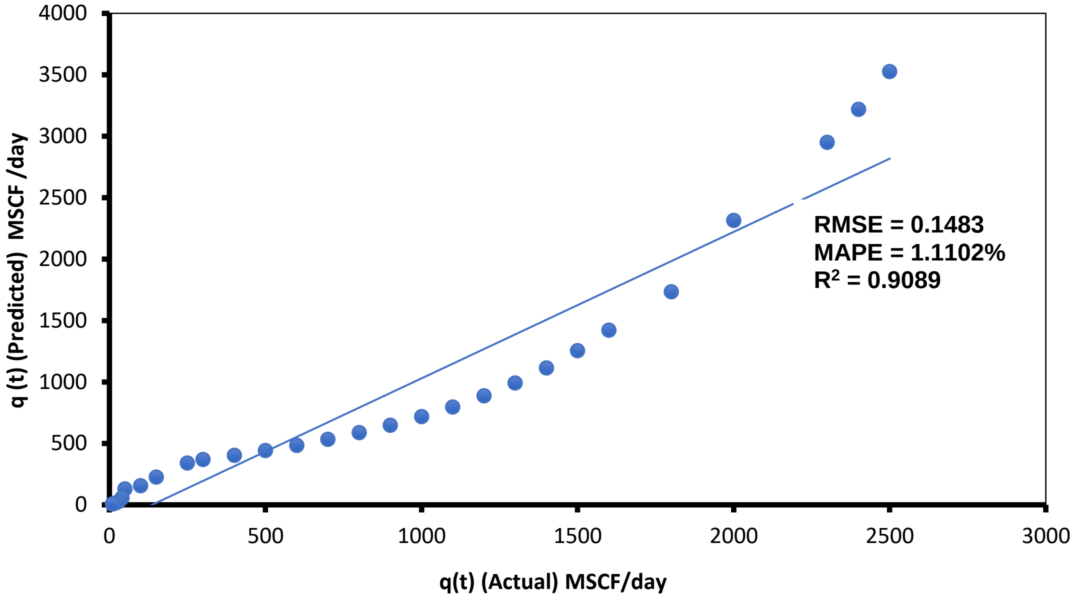
Figure 7.13: Predicted q(t) vs. actual q(t) for the Duong's decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.3. The PLE Decline Model

Figure 7.14 illustrates the scatter plot of the predicted *q(t)* values by the PLE decline model against the experimental *q(t)* values. R^2 values of 0.9710 and 0.9089 were obtained for the data sets while the MAPE of 0.5097 and 1.1102% was calculated from the data.



(a)

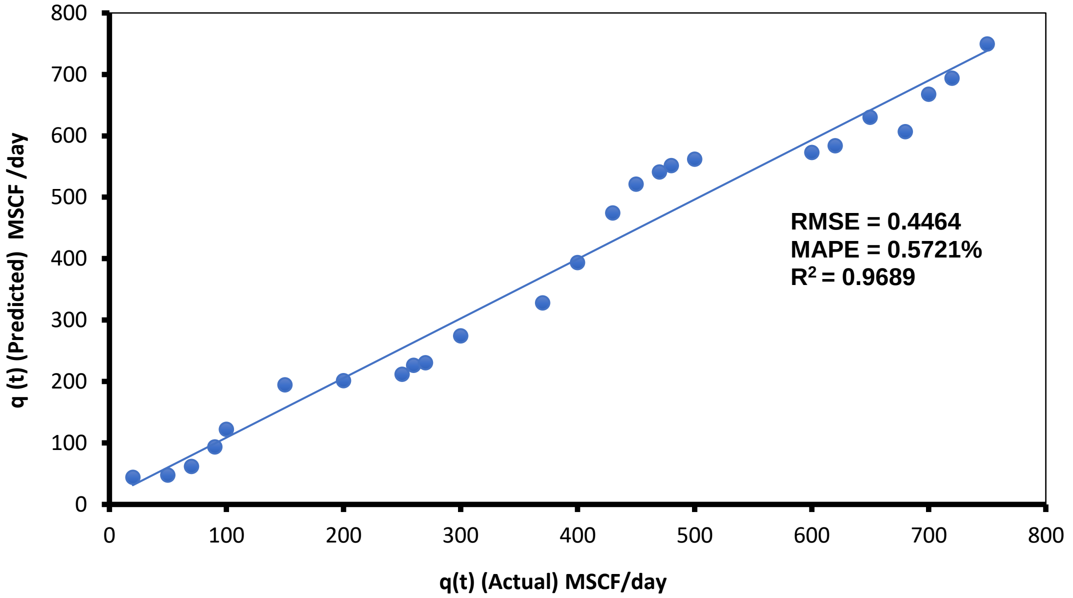


(b)

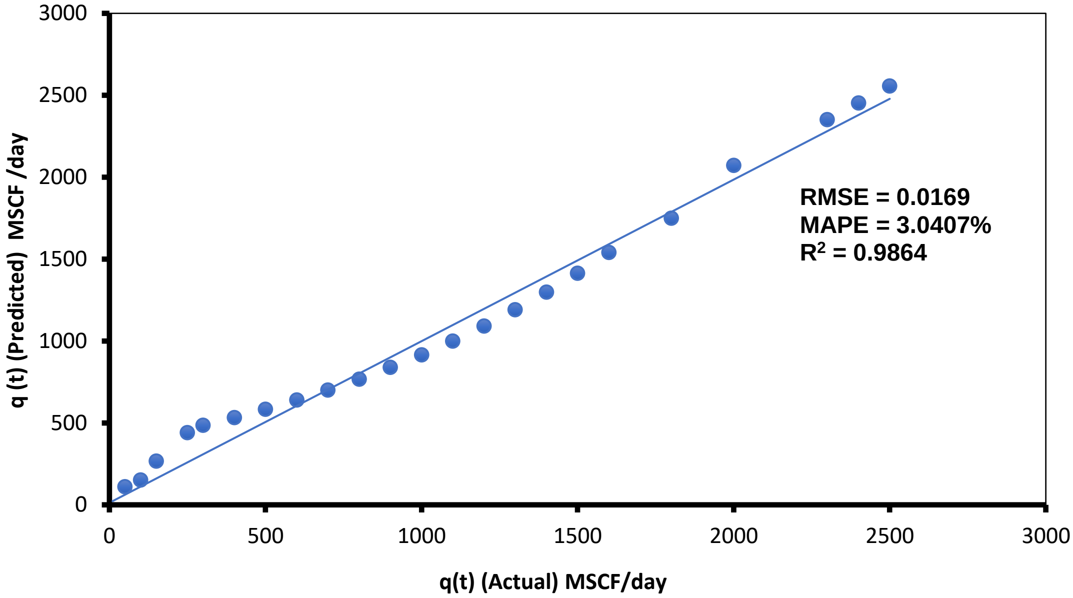
Figure 7.14: Predicted q(t) vs. actual q(t) for the PLE decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.4. The Arps-Duong-PLE Hybrid Decline Model

Figure 7.15 illustrates the scatter plot of the predicted q(t) values by the Arps-Duong-PLE hybrid decline model against the experimental q(t) values. R^2 values of 0.9689 and 0.9864 were obtained for the data sets, while the MAPE of 0.5721 and 3.0407% respectively, was calculated from the data.



(a)

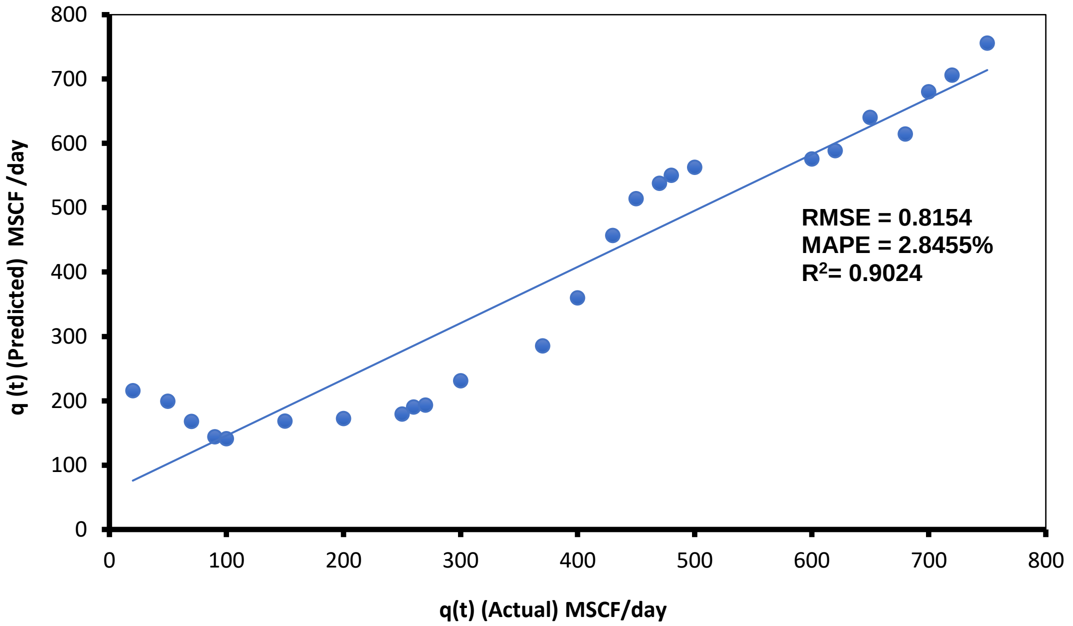


(b)

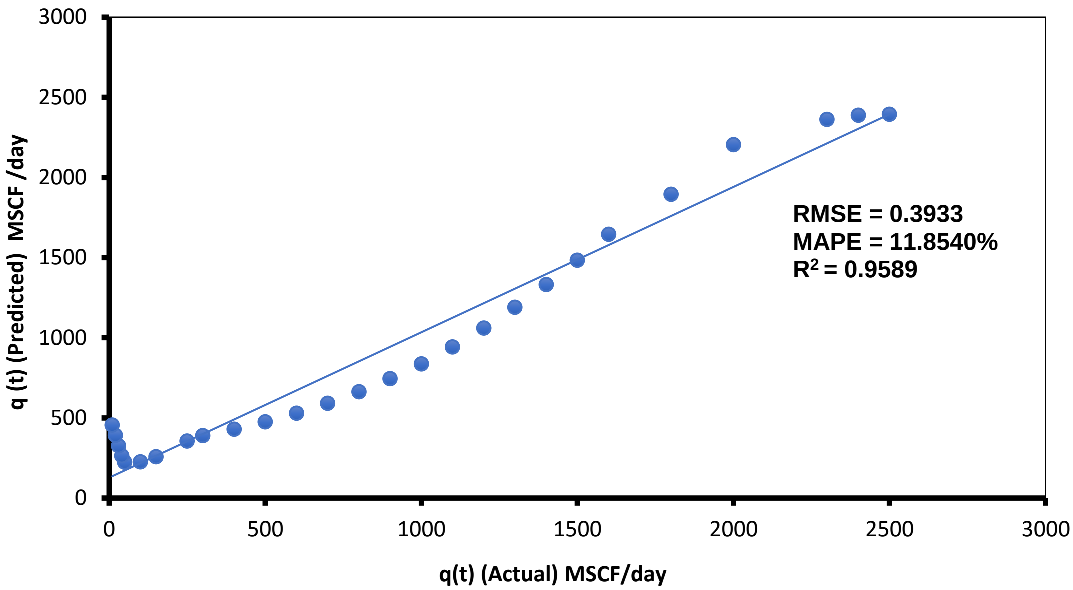
Figure 7.15: Predicted q(t) vs. actual q(t) for the Arps-Duong-PLE hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.5. The Arps-Duong Hybrid Decline Model

Figure 7.16 illustrates the scatter plot of the predicted *q(t)* values by the Arps-Duong hybrid decline model, against the experimental *q(t)* values. R^2 values of 0.9024 and 0.9589 were obtained for the data sets while the MAPE of 2.8455 and 11.8540% was calculated from the data.



(a)

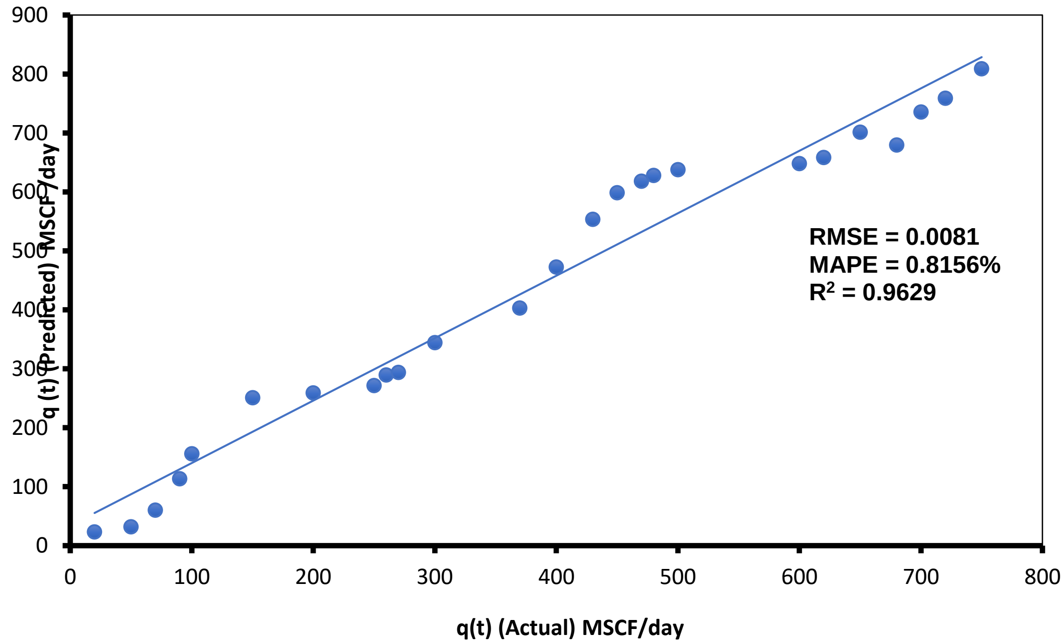


(b)

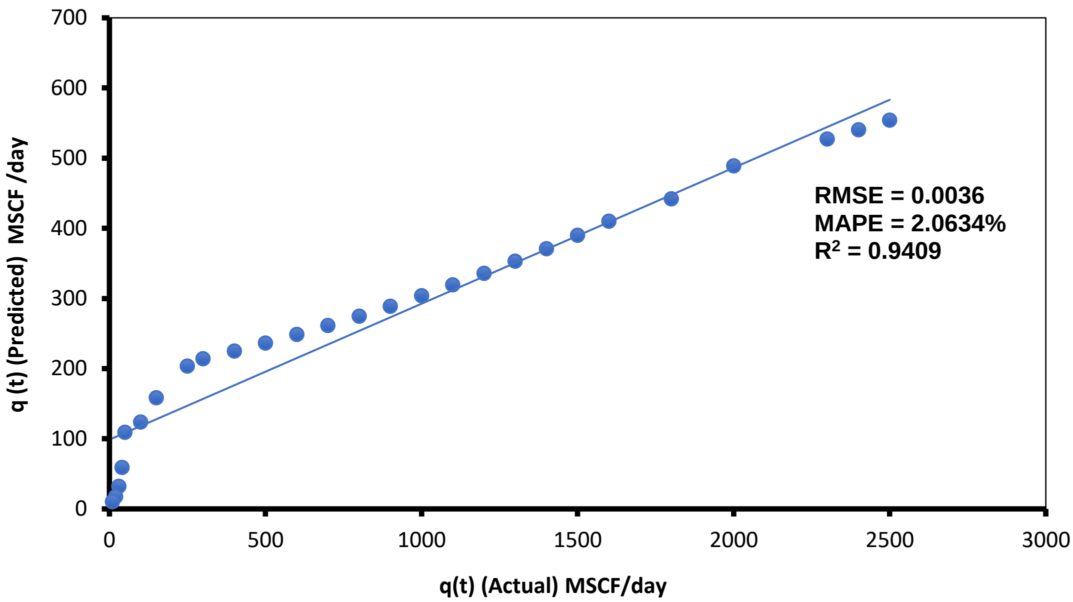
Figure 7.16: Predicted q(t) vs. actual q(t) for the Arps-Duong hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.6. The Arps-PLE Hybrid Decline Model

Figure 7.17 illustrates the scatter plot of the predicted *q(t)* values by the Arps-PLE hybrid decline model against the experimental *q(t)* values. R^2 values of 0.9629 and 0.9409 were obtained for the data sets while the MAPE of 0.8156 and 2.0634% was calculated from the data.



(a)

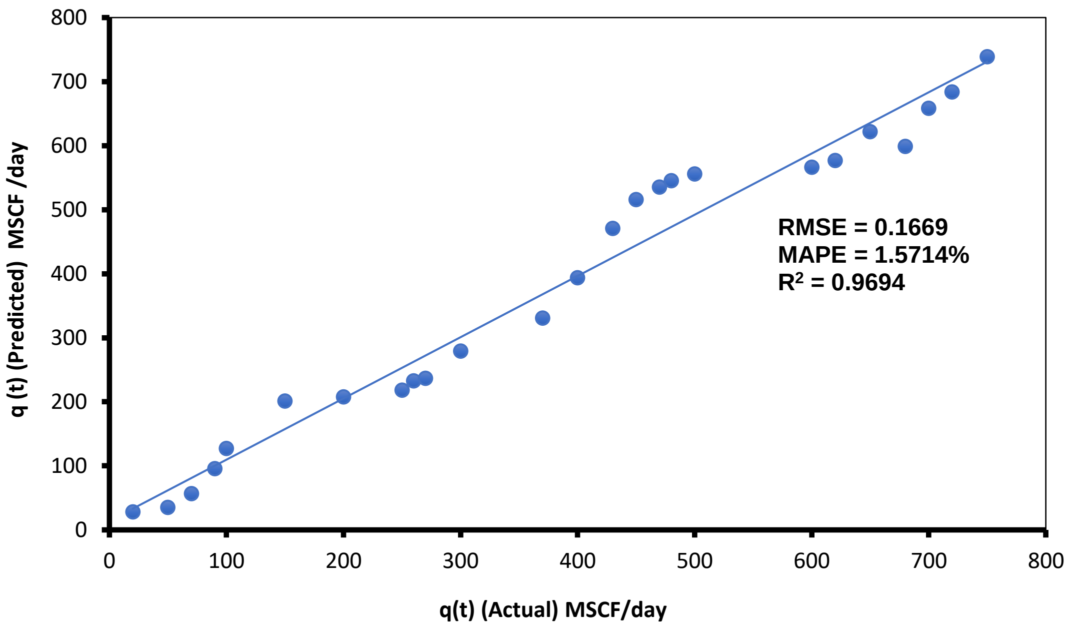


(b)

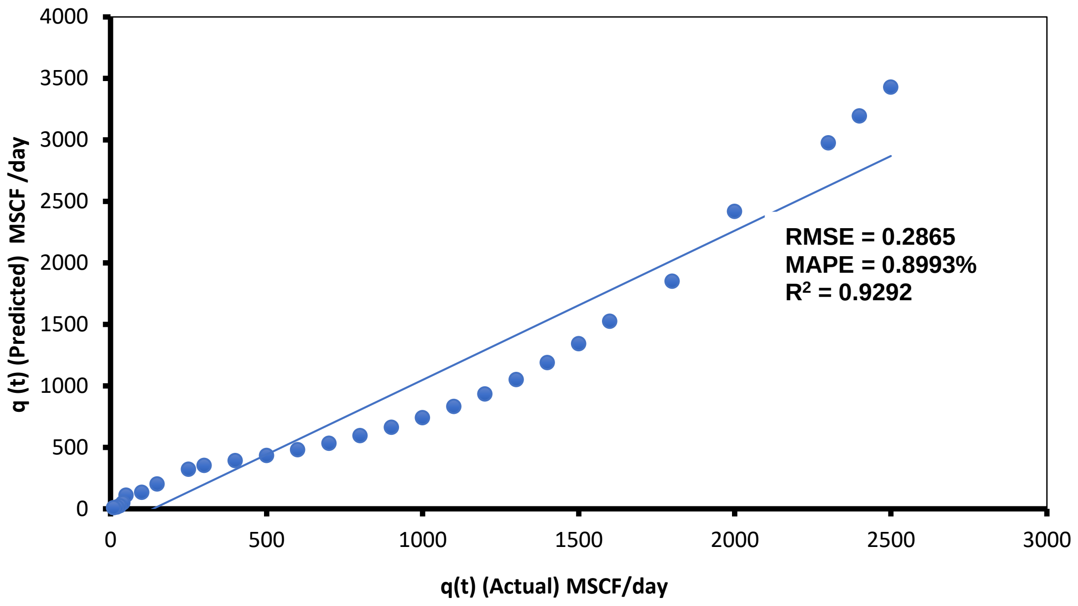
Figure 7.17: Predicted q(t) vs. actual q(t) for the Arps-PLE hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.7. The Duong-PLE Hybrid Decline Model

Figure 7.18 illustrates the scatter plot of the predicted *q(t)* values by the Duong-PLE hybrid model_against the experimental *q(t)* values. R^2 values of 0.9694 and 0.9292 were obtained for the data sets while the MAPE of 1.5714 and 0.8993% was calculated from the data.



(a)

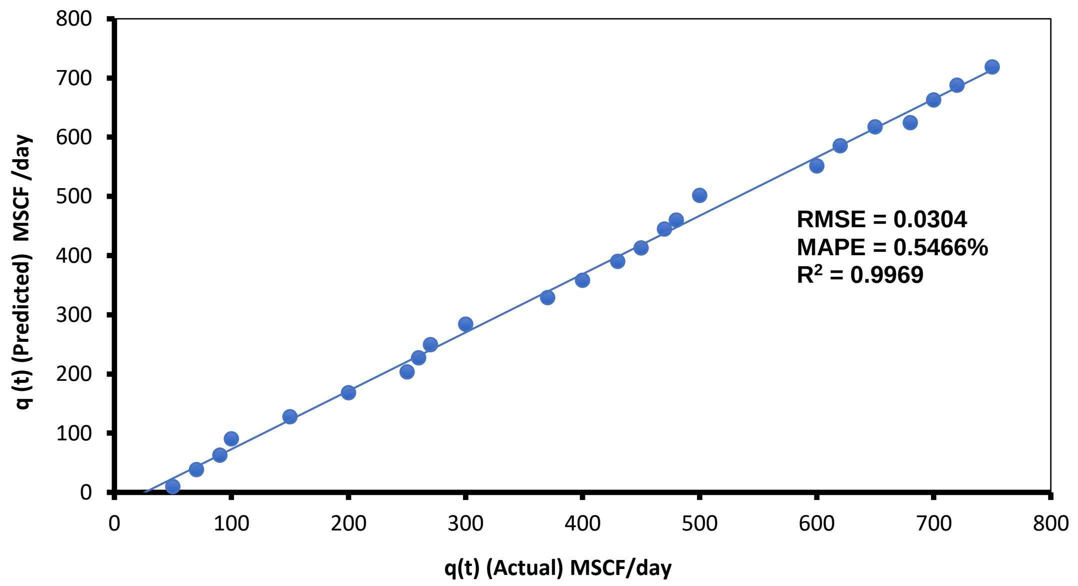


(b)

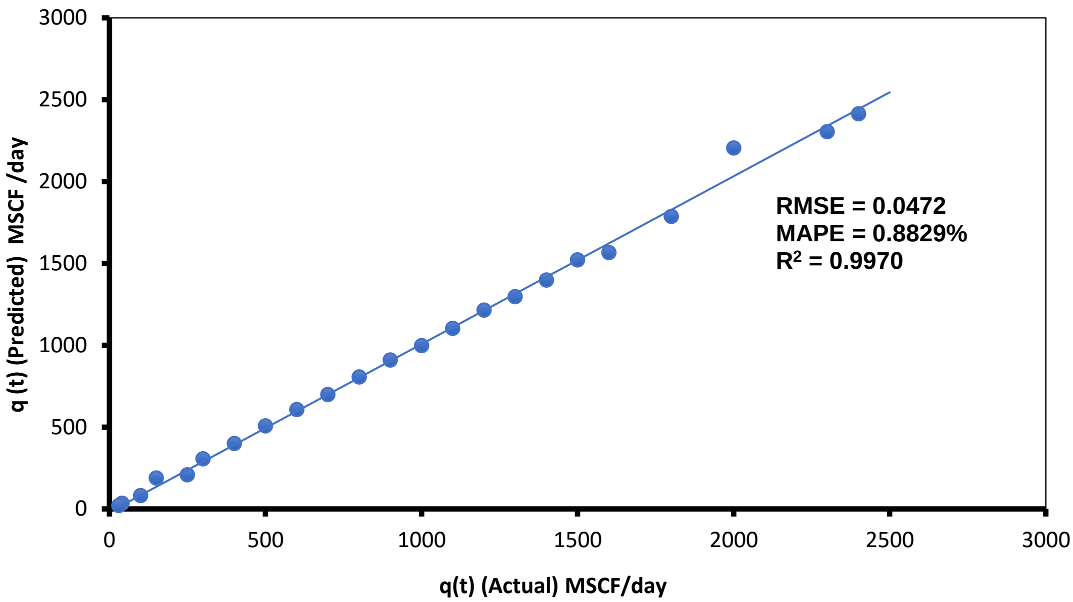
Figure 7.18: Predicted q(t) vs. actual q(t) for the Duong-PLE hybrid decline model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.8. The ARIMA Model

Figure 7.19 illustrates the scatter plot of the predicted *q(t)* values by the ARIMA model against the experimental *q(t)* values. f R^2 values of 0.9969 and 0.9970 were obtained for the data sets, while the MAPE of 0.5466 and 0.8829% was calculated from the data.



(a)

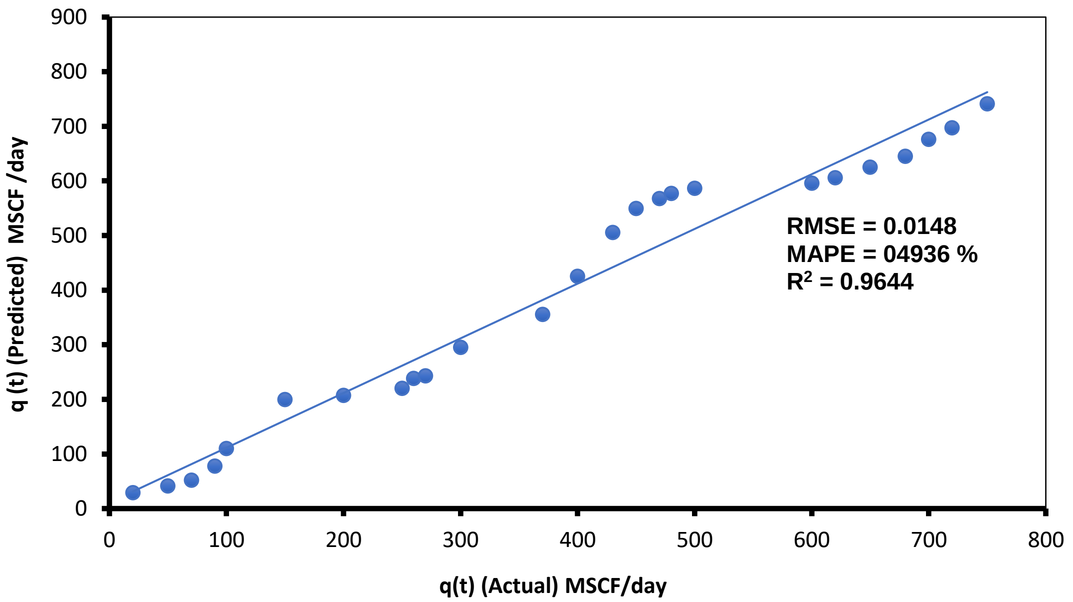


(b)

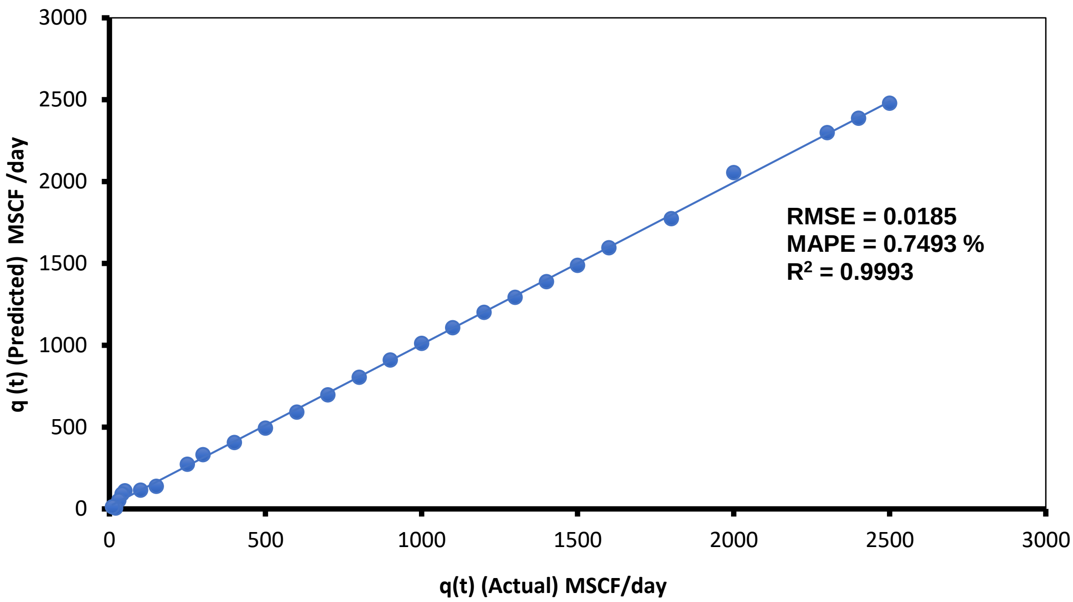
Figure 7.19: Predicted q(t) vs. actual q(t) for the ARIMA model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.9. The ANN Model

Figure 7.20 illustrates the scatter plot of the predicted *q(t)* values by the ANN model against the experimental *q(t)* values for the Canon Shale Well and Marcellus Shale Well. R^2 values of 0.9644 and 0.9993 were obtained for the data sets while the MAPE of 0.4936 and 0.7493% was calculated from the data.



(a)

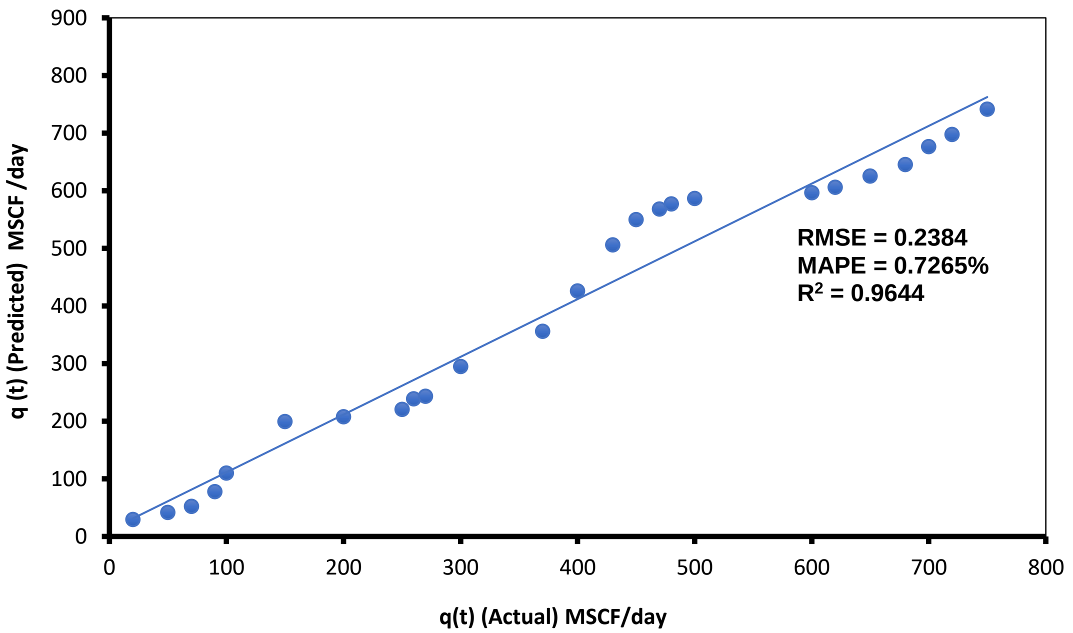


(b)

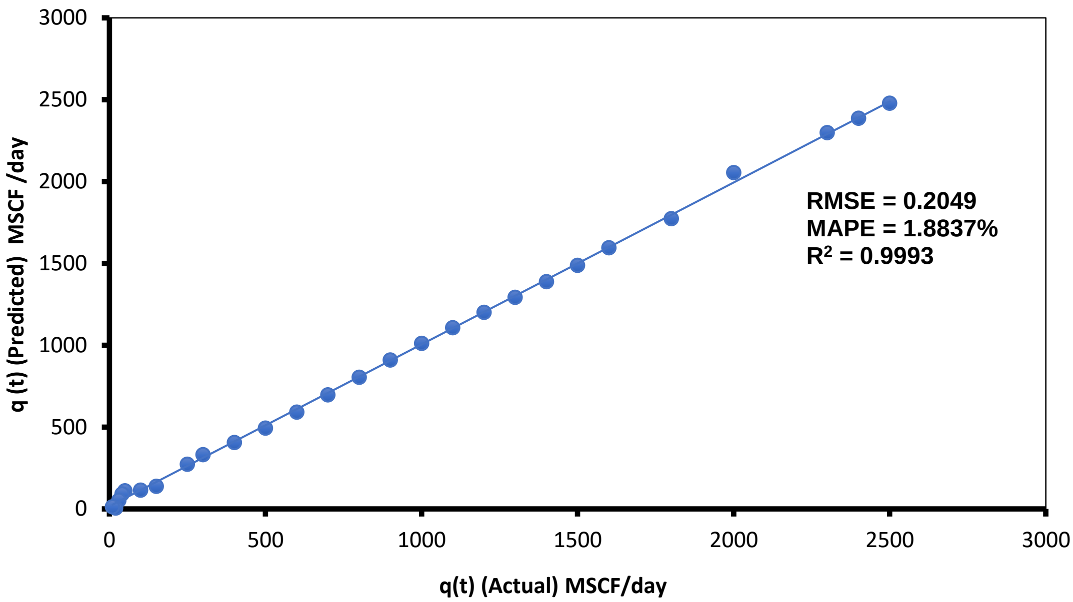
Figure 7.20: Predicted q(t) vs. actual q(t) for the ANN model for (a) Canon Shale Well and (b) Marcellus Shale Well

7.5.10. The ANN-ARIMA Hybrid Model

Figure 7.21 illustrates the scatter plot of the predicted *q(t)* values by the ANN-ARIMA hybrid model_against the experimental *q(t)* values. R^2 values of 0.9644 and 0.9993 were obtained for the data sets while the MAPE of 0.7265 and 1.8837% was calculated from the data.



(a)



(b)

Figure 7.21: Predicted q(t) vs. actual q(t) for the ANN-ARIMA hybrid model for (a) Canon Shale Well and (b) Marcellus Shale Well

Table 7.2: Summary of the accuracy and validation results for each model using the RMSE, MAPE and R^2

	Canon Shale Well			Marcellus Shale Well		
	RMSE	**MAPE (%)**	**R^2**	**RMSE**	**MAPE (%)**	**R^2**
Arps-PLE Hybrid Model	0.0081	0.8156	0.9629	0.0036	2.0634	0.9409
ANN Model	0.0148	0.4936	0.9644	0.0185	0.7493	0.9993
ARIMA Model	0.0304	0.5466	0.9969	0.0472	0.8829	0.9970
Duong's Decline Model	0.0627	0.7737	0.9388	0.2261	1.8079	0.7355
PLE Decline Model	0.0716	0.5097	0.9710	0.1483	1.1102	0.9089
Arps Decline Model	0.0725	0.5136	0.9695	0.1279	0.923	0.9447
Duong-PLE Hybrid Models	0.1669	1.5714	0.9694	0.2865	0.8993	0.9292
ANN-ARIMA Hybrid Model	0.2304	0.7265	0.9644	0.2049	1.8837	0.9993
Arps-Duong's-PLE Hybrid Model	0.4464	0.5721	0.9689	0.0169	3.0407	0.9864
Arps-Duong Hybrid Model	0.8154	2.8455	0.9024	0.3933	11.854	0.9589

Based on the summary results presented in **Table 7.2**, which ranks the models based on the RMSE value from smallest to highest, it was seen that the smaller the RMSE value, the better the model. The RMSE value was used because it is more accurate than the MAPE value. The only difference between the RMSE and MAPE values, is that the MAPE measures the deviation from the actual data in terms of percentage.

It is evident that the Arps-PLE hybrid decline model is a better model than the other models, followed by the ANN and ARIMA models, based on the RMSE value. It is apparent by combining the Arps and PLE decline models that the limitations of the models are reduced.

There is a significant reduction in RMSE values from 0.0716 and 0.0725 for the PLE and Arps models, respectively to 0.0081 for the Arps-PLE hybrid decline model for the Canon Shale Well; while for the Marcellus Shale Well, 0.1483 and 0.1279 to 0.0036. The contributing factor that was highlighted earlier is the dominance of the PLE parameters in the model i.e., D_i and D_∞ which considers the TFR flows not catered for in the Arps decline model. In addition, the PLE model was specifically developed for SGRs and by combining the models, the number of variables are reduced from four to three. This reduces the number of variables to solve, which was identified as a limitation of the model and hence the degrees of freedom is reduced.

When all the validation parameters RMSE, MAPE and R^2 were considered, the results showed that the neural network method managed to outperform the other methods. The finding concurs with this study conducted by Dhini et al. (2015). They attributed the accuracy to the non-linearity. However, contrary to their assumption, the ARIMA model proved to be the second most accurate model. As mentioned by Khashei and Bijari (2011), traditional methods such as Box–Jenkins model and ARIMA require the assumption that the time series data used in forecasting are linear, therefore they are not suitable for predicting data is non-linear.

The reasoning for the ARIMA working well for predicting the decline behaviour may be attributed to different flow regimes that can occur in a hydraulically fractured reservoir. The different flow regimes do follow a linear trend. It can be assumed that the weak performance of the other decline curve hybrid models is due to the changes in flow regimes; an effect of which is not considered. In addition, one would have to consider extending the data range over a longer period, to determine whether the hybrid models would yield results that are more positive.

The results also highlighted that the ANN model had an advantage over the ANN-ARIMA hybrid model, contrary to this study conducted by Zhang (2001). The ANN-ARIMA model did not fare well which contradicts the literature which may be the result of using one input. The model had a high RMSE value.

As mentioned earlier in Chapter 7, this concurs with the work conducted by Taskaya-Temizel and Ahmad (2005). According to Taskaya-Temizel and Ahmad (2005), two factors prevent the ANN-ARIMA hybrid model from delivering better results.

These are:

1. The assumption of the existence of a relationship between the components of the linear and non-linear components in the data can cause performance degradation as other model relationships (e.g., multiplicative) may exist within the data instead of linear/non-linear relationships and,
2. Secondly, there is no guarantee that the residual of the linear components will have valid non-linear patterns.

Therefore, the results show that the Arps-PLE hybrid model gave predicted values closest to the actuals. The accuracy evaluation indicated that the model is consistent in predicting production rates. Lastly, the validation process concurred with the accuracy evaluation since the model yielded a low RMSE compared to the ANN and ARIMA models.

7.6. Evaluation of the Confidence Intervals for Arps-PLE Hybrid Decline and ANN

After the validation of the models, it was found the Arps-PLE hybrid decline and ANN models to be the best models in predicting shale gas production. A step was taken further and 95% confidence intervals were assessed. **Figures 7.22** and **7.23**, illustrate the findings[14].

[14] Raw data in the Appendices – A.4. Confidence Intervals

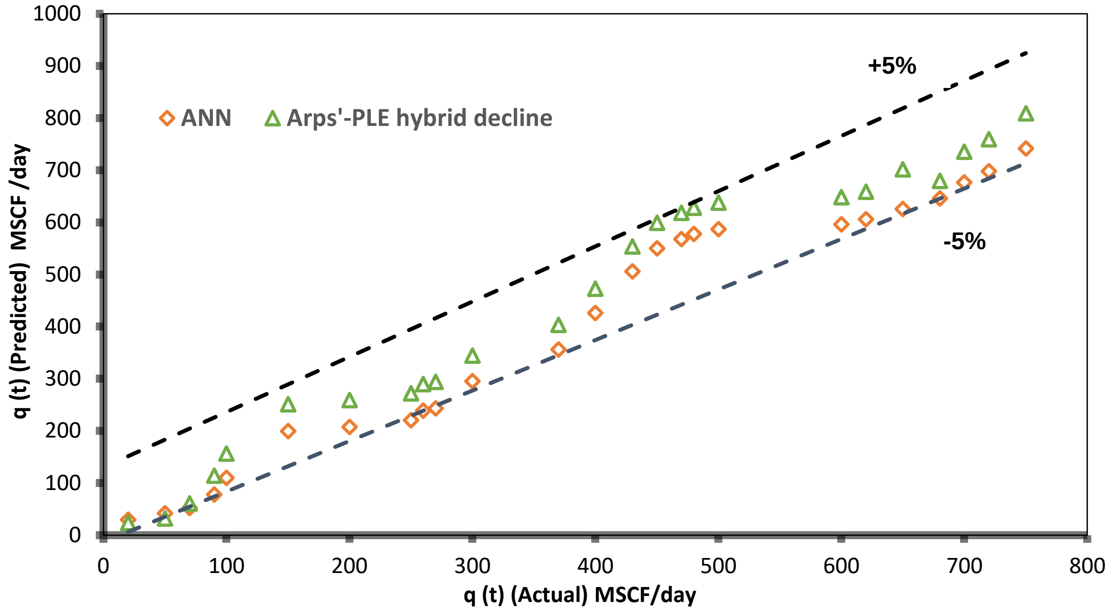


Figure 7.22: Scatter plot of the predicted shale production data by the Arps-PLE hybrid decline and ANN models vs. actual data for the Canon Shale Well evaluating the 95% confidence level

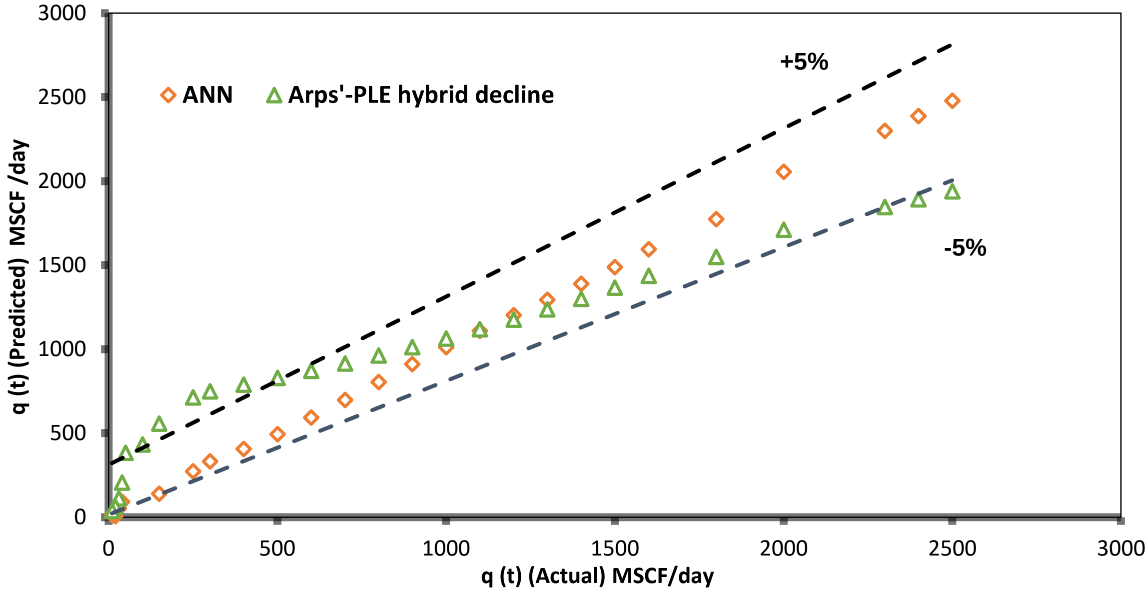


Figure 7.23: Figure 7: Scatter plot of the predicted shale production data by the Arps-PLE hybrid decline and ANN models vs. actual data for Marcellus Shale Well evaluating the 95% confidence level

Evaluating the confidence intervals is one way to help assess what the values might be in the wider population (Fethney, 2010). Confidence intervals provide the possible range of values, bracketed by lower and upper limits that encompass the unknown population or ‘true’ value estimated by that sample mean, correlation coefficient or odds ratio (Fethney, 2010). It is usual to report either the 90%, 95% or 99% confidence interval (CI); the 95% CI tends to be commonly used (Fethney, 2010). From **Figures 7.22 and 7.23**, it is evident that for the Arps-PLE hybrid decline and ANN model there is a confidence that 95% of the data lies within the ±5% error band. The results demonstrate very good fitting between the actual and predicted values for the models.

7.7. Conclusion

The objective of this chapter was to evaluate the forecasting performance of decline curve hybrid models and ANN-ARIMA hybrid model with Arps; Duong's; PLE decline models; ARIMA and ANN models, respectively. The experimental results were obtained using the different prediction models i.e., Arps, Duong's, PLE, Arps-Duong's-PLE hybrid decline, Arps-Duong's hybrid decline, Arps-PLE hybrid decline, Duong's-PLE hybrid decline, ARIMA, ANN and lastly the hybrid ANN-ARIMA model.

The following can be concluded:

- The current DCA models, Arps, Duong's and PLE decline models appear to over and underestimate the data.
- The DCA hybrid models, Arps-Duong's-PLE hybrid decline, Arps-Duong's hybrid decline, Duong's-PLE hybrid decline did not give the best outcome as assumed it would, in comparison to the individual DCA models. However, the Arps-PLE Hybrid decline model gave the closest predicted results.
- Both the ARIMA and ANN models gave good, predicted results. However, when both models were combined into the ANN-ARIMA hybrid model the strengths of both models referenced in literature did not provide accurate predictive data, which can be attributed to the use of a single input. The resultant was an overestimation in the production flow rate.
- Overall, the models, which gave predicted values closest to the actuals was the ARIMA, ANN and the Arps- PLE hybrid decline models. The ANN-ARIMA maybe included if the input for the ANN is increased.
- In the model accuracy evaluation, the Arps-PLE hybrid decline, ARIMA and ANN gave consistent prediction results between the two sets of data evaluated (Well #A and Barnett Shale Well).

- The validation of the models indicated that the Arps-PLE hybrid decline model gave the lowest RMSE value with a good R^2 value for both the Canon Shale Well and Marcellus Shale Well, followed by the ANN and ARIMA models.
- The weak performance of the other decline curve hybrid models is caused by; either changes in flow regimes that are not considered or the shorter period within which the data is used.
- Lastly, the confidence interval evaluation found that the Arps-PLE hybrid decline, and ANN model fell within the 95% confidence limit, i.e., the data lies within the ±5% error band. The results demonstrate very good fit between the actual and predicted values for the models. This corresponds with the findings from the validation process.

In conclusion, the findings have provided a significant contribution to the prediction of shale gas production. The results indicate that the Arps-PLE hybrid decline model is a good model predictor for shale gas production. A contributing factor is the dominance of the PLE parameters i.e., D_i changes at early stages and D_∞ becomes constant at a later time in the model. This caters for the TFR which the Arps decline model did not consider. Lastly, with the PLE model, the limitation identified during the sensitivity analysis was the number of variables. Therefore, by combining the models the number of variables is reduced and the degrees of freedom is reduced from four to three.

Chapter 8

Financial Model Analysis of Shale Gas

8.1. Introduction

Shale gas production is a profitable business for oil and gas operators, provided there is assurance that gas can be produced commercially and in a sustainable way (Guarnone et al., 2012). The present practice of shale gas economic valuation commonly uses a mean EUR and a single production decline model for the whole lease or play (Penner, 2013). Nevertheless, shale gas production remains economically risky because the EUR remains poorly constrained during the early stages of field development (Weijermars, 2013).

Gas well revenues are a function of two key variables (Lake et al., 2013):

- The price per thousand cubic feet (Mcf) of natural gas sold, the volume of gas produced, and
- The volume of gas it will produce.

Guarnone et al. (2012) pointed out that cost estimate plays a key role because it supports the economic evaluation and support process for buying an exploration permit or not. The cost structure of a shale gas project differs from conventional production, which makes cost estimation problematic. Guarnone et al. (2012) also mentioned that appraising noteworthy OPEX-like investments involves an accurate prediction of the time when they will be required, which in turn depends on the prediction of DCA to model single wells and overall field performance.

Since EUR calculated from DCA models plays a significant role in the economic analysis of shale gas development, the accuracy thereof is crucial for exploration. In this chapter, the ANN and Arps-PLE hybrid models discussed in the preceding chapters will be used to evaluate the economic viability of shale gas reservoirs.

The ANN and Arps-PLE hybrid decline models, which were found to predict values closest to the actuals, were used to calculate the EUR, compared with other DCA models. The philosophy and methodology by Lake et al. (2013) in their study will be used. Their study indicated that the production estimates are based on a combination of two types of production decline curves i.e., initially a hyperbolic decline and then switching to an exponential decline.

8.2. Well Productivity using Estimated Ultimate Recovery (EUR)

According to Weijermars (2013), understanding well productivity of shale gas plays provides important guidance for the economic development of shale gas wells (Weijermars, 2013). In chapter 7, the ANN and Arps-PLE hybrid decline models were evaluated and validated and it was found that the two models delivered good forecasting results. Hence, the predicted data from the models will be used to estimate the EUR and ultimately evaluate the economic analysis.

Assuming the philosophy by Lake et al. (2013), **Figure 8.1 (a), (b)** and **(c)** shows the hyperbolic decline for the predicted data from ANN while **Table 8.1** summarizes the q_i and D_i values.

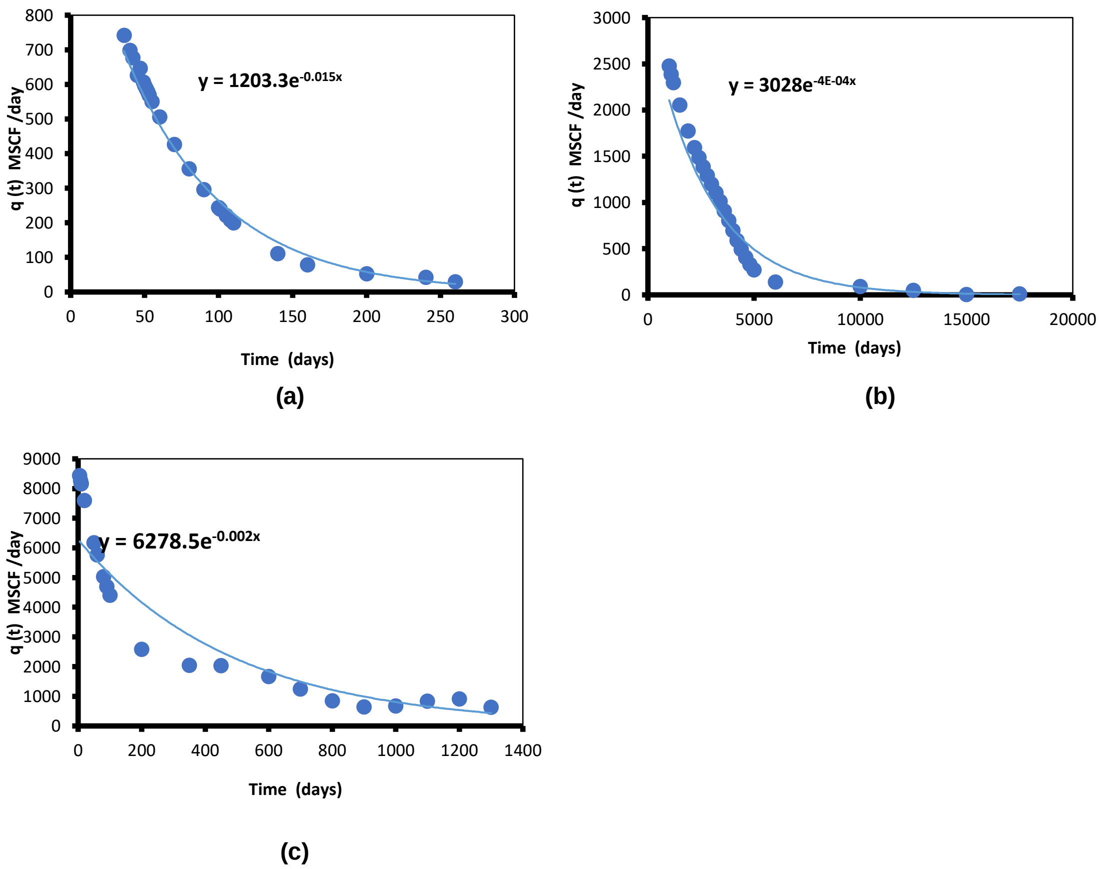


Figure 8.1: Predicted q(t) vs. time for the ANN model for (a) Canon Shale Well, (b) Marcellus Shale Well and Barnett Shale Well

Table 8.1: Summaries the q_i and D_i values for each of the data sets from the various shale plays for the ANN model

	q_i	D_i
Cannon Well	1203	-0.015
Marcellus Shale Well	3028	-0.0004
Barnett Shale Well	6279	-0.002

Figure 8.2 (a), (b) and **(c)** shows the hyperbolic decline for the predicted data from the Arps-PLE hybrid decline model, while **Table 8.2** summarizes the q_i and D_i values.

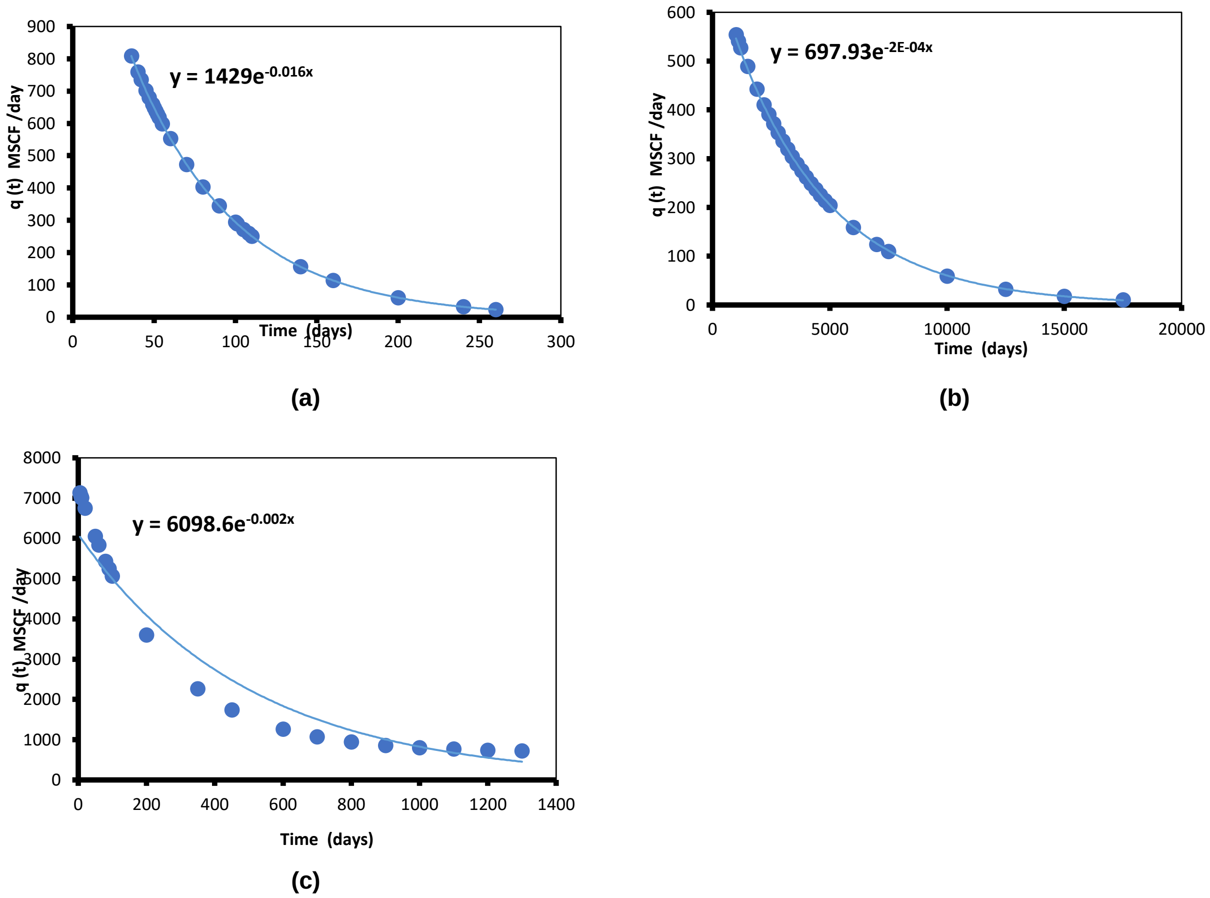


Figure 8.2: Predicted q(t) vs. time for the Arps-PLE hybrid decline model for (a) Canon Shale Well, (b) Marcellus Shale Well and Barnett Shale Well

Table 8.2: Summaries the q_i and D_i values for each of the data sets from the various shale plays for the Arps-PLE hybrid decline model

	q_i	D_i
Cannon Well	1429	-0.016
Marcellus Shale Well	698	-0.0004
Barnett Shale Well	6099	-0.002

The production estimates for the Cannon Well, the Marcellus Shale Well and Barnett Shale Well was based on a combination of two types of production decline curves. Initial production declines are based on a hyperbolic decline curve, as shown in **Figures 8.1** and **8.2,** respectively.

The production decline curve switches to an exponential decline when the rate of decline is 1.5 and 1.6% for the Cannon Well, 0.04% for Marcellus Shale Well and 0.2% for the Barnett Shale Well, respectively. This production decline function used to model is a hybrid model approach, containing a mixture of hyperbolic and exponential functions.

According to Lake et al. (2013), given the relatively brief time that horizontal, hydraulically fractured gas wells have been producing, the production life of these wells is relatively still unknown. They went on to explain that although conventional wells have been producing for 30 years and longer.

The initial evidence from Barnett Shale Well and Marcellus Shale Well indicate short lifespans i.e., for the Barnett Shale Well the average life of a well is 7.5 years (Lake et al., 2013).

The short lifespan does not in any way mean that the wells are not economic. However, to be profitable; the volume of the initial production has to be lucrative to fund the investment.

The switch from the hyperbolic to the exponential decline function occurs between days (indicated by when the graphs appear to become flatter):

- 200 - 250 days (Cannon Well)
- 10000 - 12000 days (Marcellus Shale Well)
- 800 – 1000 days (Barnett Shale Well)

The switching between the two is important because it provides the EUR for the well. **Figures 8.1 and 8.2** shows the estimated production volumes. The figures show the initial steep decline in the production rate.

Using the estimated production data, the cumulated production (N_p) was calculated using Equation (8.2).

$$N_p = \frac{q_i - q_t}{D} \tag{8.2}$$

Figure 8.3 shows the N_p over the period for the (a) Canon Shale Well, (b) Marcellus Shale Well and (c) Barnett Shale Well for the ANN model.

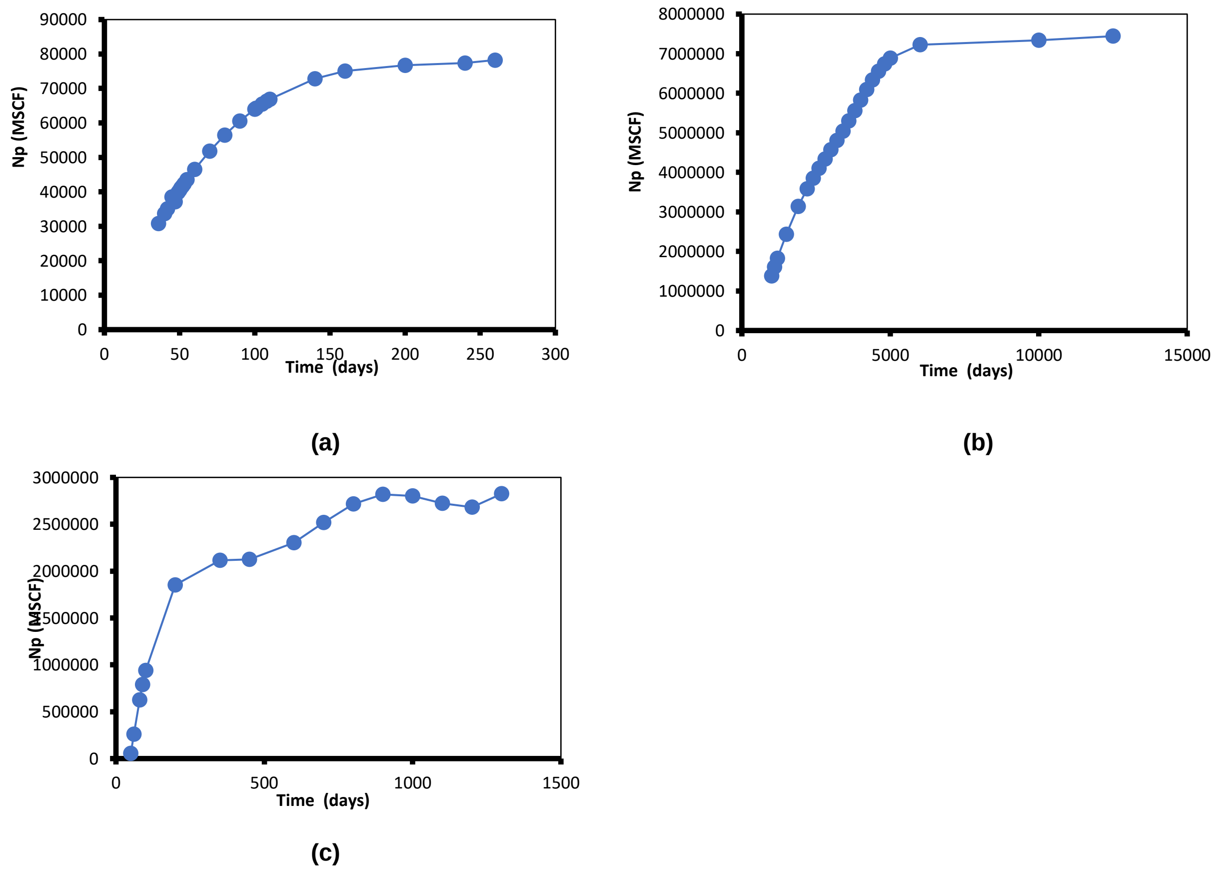
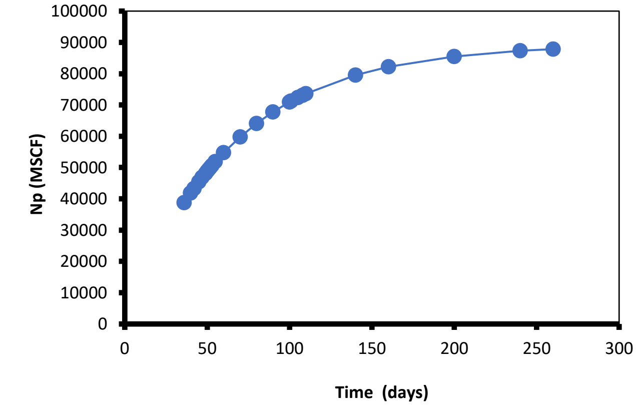
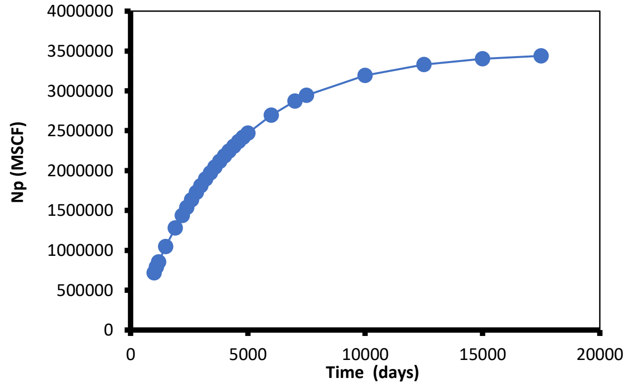


Figure 8.3: Cumulative production (Np) vs. time for (a) Canon Shale Well, (b) Marcellus Shale Well and (c) Barnett Shale Well for the ANN model

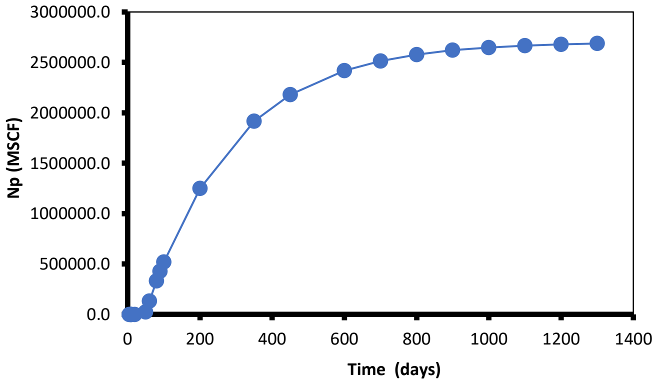
Figure 8.4 shows the N_p over the period for the (a) Canon Shale Well, (b) Marcellus Shale Well and (c) Barnett Shale Well for the Arps-PLE hybrid decline model.





(a)

(b)



(c)

Figure 8.4: Cumulative production (Np) vs. time for (a) Canon Shale Well, (b) Marcellus Shale Well and (c) Barnett Shale Well for the Arps-PLE hybrid decline model

Cumulative production is used to determine the economic limit of the wells potential i.e., the EUR. As previously done, a comparison was done between the use of a hyperbolic-exponential model to predict the depletion of the natural producing well. According to Makinde et al. (2012), the issue with this approach in most cases is that the hyperbolic over-estimates while the exponential model under-estimates. However, the predicted results have satisfactorily validated the reasoning for using this approach in this study.

The EUR estimated for the period under review obtained from **Figures 8.3 and 8.4** were:

- 0.079 and 0.087 bcf (Canon Shale Well)
- 7.4 and 3.5 bcf (Marcellus Shale Well)
- 2.8 and 2.6 bcf (Barnett Shale Well)

As mentioned, to understand the well productivity that representative US shale gas plays provides, important guidance for the economic development of shale gas wells in emergent shale plays elsewhere in the world needs to be provided (Weijermars, 2013). A review of the US well productivities, using 46,506 shale gas wells, gives a 40 year mean EUR of 1.14 bcf (Weijermars, 2013). A study conducted by Swindell (2018) which investigated the EUR of over 5000 wells in the Marcellus Shale Well, stated that their results indicated that each well had a mean of 5.0 bcf. However, the results varied between 7.2 to 12 bcf per county (Swindell, 2018). Moeller and Murphy, 2016 found that the range for the EUR for the range of Marcellus wells was between 0.23 and 6.0 bcf with an average of 3.2 bcf.

While Weijermars (2013) mentioned that for the Barnett Shale Well, the best areas have a mean EUR of 2.1 bcf/well and the worst areas have a mean of 0.59 bcf/well. Contrary to Weijermars (2013), Patzek et al. (2013), found that the lower bounds of the Barnett had a EUR average of 1.0 bcf while the upper bounds were found to have an average of 7.0 bcf. Browning and Tinker (2013) obtained EUR ranging from 0.4 to 4.3 bcf for the Barnett. Canon Shale Well and indicated a low EUR of 0.079 and 0.087 bcf, respectively, The literature showed an average value of 0.0013 bcf (226 000 bbl) among the 975 wells (Oils and Gas Journal, 2020).

As detailed in Chapter 6, It is important to highlight the Evaluation and Sensitivity of Decline Curve Models. The accuracy assessment indicated that decline curve modelling impacts the EUR of SGRs, and it was observed that all decline models yield a different EUR result, which is either is over or under-estimated. Studies have revealed that the production time significantly impacts the EUR, depending on which decline model is being used (Manda and Nkazi, 2020). However, in Chapter 7, The Forecasting, Accuracy and Validation of Hybrid Models; it was found that the ANN Model aligned very well with the actual production data. **Table 8.3** compares the EUR values between the ANN, Arps-PLE hybrid decline models and other decline models. The table shows the different EUR values obtained from the different models. During the evaluation of the models, it was found that the Arps, Duong's and PLE models over and under-estimated production data, which would ultimately result in either an over or under-estimation of the EUR.

Table 8.3: Comparison of EUR values from the different models

	Arps Decline Model	**Duong's Decline Model**	**PLE Model Decline Model**	**ANN Model**	**Arps-PLE hybrid decline Model**
Canon Shale Well (bcf)	0.100	0.152	0.101	0.079	0.087
Marcellus Shale Well (bcf)	9.24	16.14	9.09	7.40	3.50
Barnett Shale Well (bcf)	5.11	5.03	2.04	3.20	2.60

8.3. Well Economics

Tables 8.4, 8.5, 8.6, 8.7, 8.8 and 8.9 provides a summary of the financial estimates for the Canon Shale Well, Marcellus Shale Well, and lastly, the Barnett Shale Well for two models, respectively. The estimated gross gas revenue can be calculated by multiplying the annual production volumes by the price/Mcf for natural gas. The adjustment of these estimates for the working interest of the investor produces a net gas revenue. The annual cash flows are calculated using the equation (2.27).

Table 8.4: Financial estimates for the Canon Shale Well, using the ANN model

Time (day)	Estimated Gross Gas (Mcf)	Estimated Net Gas (15.75% of Gross Gas) (Mcf)	Price of Natural Gas /Mcf ($)	Estimated Total Net Revenue ($)[15]
36	30787.68	4849.06	2.77	13,431.89
40	33702.56	5308.15	2.77	14,703.59
42	35114.63	5530.55	2.77	15,319.64
45	38513.94	6065.95	2.77	16,802.67
47	37176.53	5855.30	2.77	16,219.19
49	39821.83	6271.94	2.77	17,373.27
50	40464.77	6373.20	2.77	17,653.77
51	41100.41	6473.32	2.77	17,931.08
52	41728.78	6572.28	2.77	18,205.22
53	42349.91	6670.11	2.77	18,476.21
55	43570.55	6862.36	2.77	19,008.74
60	46497.44	7323.35	2.77	20,285.67
70	51830.38	8163.28	2.77	22,612.30
80	56497.95	8898.43	2.77	24,648.65
90	60538.46	9534.81	2.77	26,411.42
100	63993.51	10078.98	2.77	27,918.77
101	64308.50	10128.59	2.77	28,056.19
105	65515.26	10318.65	2.77	28,582.67
108	66365.74	10452.60	2.77	28,953.72
110	66907.44	10537.92	2.77	29,190.05
140	72873.90	11477.64	2.77	31,793.06
160	75027.98	11816.91	2.77	32,732.83
200	76727.16	12084.53	2.77	33,474.14
240	77430.75	12195.34	2.77	33,781.10
260	78258.43	12325.70	2.77	34,142.20

[15] excludes depletion allowance, total net operating expenses and lastly the severance and ad valorem expenses

Table 8.5: Financial estimates for the Barnett Shale Well using the ANN model

Time (day)	Estimated Gross Gas (Mcf)	Estimated Net Gas (15.75% of Gross Gas) (Mcf)	Price of Natural Gas /Mcf ($)	Estimated Total Net Revenue ($)
50	54019.05	8508.00	2.77	23,567.16
60	259914.03	40936.46	2.77	113,394.00
80	626945.12	98743.86	2.77	273,520.48
90	789435.39	124336.07	2.77	344,410.92
100	938800.27	147861.04	2.77	409,575.09
200	1850359.98	291431.70	2.77	807,265.80
350	2115901.95	333254.56	2.77	923,115.12
450	2125893.21	334828.18	2.77	927,474.06
600	2304156.93	362904.72	2.77	1,005,246.07
700	2519207.51	396775.18	2.77	1,099,067.26
800	2715598.01	427706.69	2.77	1,184,747.52
900	2817458.64	443749.74	2.77	1,229,186.77
1000	2803435.33	441541.06	2.77	1,223,068.75
1100	2722007.80	428716.23	2.77	1,187,543.95
1200	2681572.93	422347.74	2.77	1,169,903.23
1300	2824253.08	444819.86	2.77	1,232,151.01

Table 8.6: Financial estimates for the Canon Shale Well using the Arps-PLE hybrid decline model

Time (day)	Estimated Gross Gas (Mcf)	Estimated Net Gas (15.75% of Gross Gas) (Mcf)	Price of Natural Gas /Mcf ($)	Estimated Total Net Revenue ($)
36	38757.85	6104.36	2.77	16,909.08
40	41855.60	6592.26	2.77	18,260.55
42	43332.55	6824.88	2.77	18,904.91
45	45462.23	7160.30	2.77	19,834.04
47	46826.94	7375.24	2.77	20,429.42
49	48149.17	7583.49	2.77	21,006.28
50	48794.77	7685.18	2.77	21,287.94
51	49430.25	7785.26	2.77	21,565.18
52	50055.76	7883.78	2.77	21,838.08
53	50671.46	7980.76	2.77	22,106.69
55	51874.04	8170.16	2.77	22,631.35
60	54719.29	8618.29	2.77	23,872.66
70	59777.53	9414.96	2.77	26,079.44
80	64096.15	10095.14	2.77	27,963.55
90	67783.29	10675.87	2.77	29,572.16
100	70931.30	11171.68	2.77	30,945.55
101	71219.59	11217.09	2.77	31,071.33
105	72328.24	11391.70	2.77	31,555.00
108	73114.91	11515.60	2.77	31,898.21
110	73619.01	11594.99	2.77	32,118.13
140	79545.59	12528.43	2.77	34,703.75
160	82193.00	12945.40	2.77	35,858.75
200	85529.52	13470.90	2.77	37,314.39
240	87302.39	13750.13	2.77	38,087.85
260	87847.24	13835.94	2.77	38,325.55

Table 8.7: Financial estimates for the Marcellus Shale Well using the Arps-PLE hybrid decline model

Time (day)	Estimated Gross Gas (Mcf)	Estimated Net Gas (15.75% of Gross Gas) (Mcf)	Price of Natural Gas /Mcf ($)	Estimated Total Net Revenue ($)
1000	718691	113194	2.77	313,547
1100	787294	123999	2.77	343,477
1200	854195	134536	2.77	372,664
1500	1045097	164603	2.77	455,949
1900	1278239	201323	2.77	557,664
2200	1438340	226539	2.77	627,512
2400	1538560	242323	2.77	671,235
2600	1633867	257334	2.77	712,815
2800	1724500	271609	2.77	752,356
3000	1810690	285184	2.77	789,959
3200	1892654	298093	2.77	825,718
3400	1970599	310369	2.77	859,723
3600	2044723	322044	2.77	892,062
3800	2115213	333146	2.77	922,814
4000	2182246	343704	2.77	952,059
4200	2245993	353744	2.77	979,871
4400	2306614	363292	2.77	1,006,318
4600	2364263	372371	2.77	1,031,469
4800	2419086	381006	2.77	1,055,387
5000	2471221	389217	2.77	1,078,132
6000	2695968	424615	2.77	1,176,184
7000	2870764	452145	2.77	1,252,443
7500	2943003	463523	2.77	1,283,958
10000	3194671	503161	2.77	1,393,755
12500	3328922	524305	2.77	1,452,326
15000	3400538	535585	2.77	1,483,570
17500	3438740	541602	2.77	1,500,237

Table 8.8: Financial estimates for the Barnett Shale Well using the Arps-PLE hybrid decline model

Time (day)	Estimated Gross Gas (Mcf)	Estimated Net Gas (15.75% of Gross Gas) (Mcf)	Price of Natural Gas /Mcf ($)	Estimated Total Net Revenue ($)
50	23356.36	3678.63	2.77	10,189.79
60	130572.75	20565.21	2.77	56,965.63
80	332308.80	52338.64	2.77	144,978.02
90	427164.80	67278.46	2.77	186,361.32
100	518225.87	81620.57	2.77	226,088.99
200	1250693.47	196984.22	2.77	545,646.29
350	1915939.32	301760.44	2.77	835,876.43
450	2179867.39	343329.11	2.77	951,021.65
600	2419573.63	381082.85	2.77	1,055,599.49
700	2514674.12	396061.17	2.77	1,097,089.45
800	2577895.16	406018.49	2.77	1,124,671.21
900	2619923.33	412637.92	2.77	1,143,007.05
1000	2647862.88	417038.40	2.77	1,155,196.38
1100	2666436.57	419963.76	2.77	1,163,299.61
1200	2678784.01	421908.48	2.77	1,168,686.50
1300	2686992.37	423201.30	2.77	1,172,267.59

Due to the limited information available regarding the equity investment, depletion allowance, total net operating expenses and lastly severance and ad valorem expenses, these parameters will be excluded from the scope and the financial model will only consider the estimated total net revenue. Using the current natural gas price of $2.77/Mcf (https://oilprice.com) and a working interest rate of 15.75%, extracted from the work of Lake et al. (2013), the financial estimates for Canon Shale Well, Marcellus Shale Well and Barnett Shale Well were generated.

From the financial analysis, it can be presumed that the production forecast and resulting EUR are sensitive to natural gas prices. That finding concurs with the study conducted by Browning et al. (2013). Their study examined detailed well economics and reserve forecasting. The results also indicated that the extraction of shale gas is a lucrative venture evident in the high estimated total net revenue for all three shale plays. The degree of influence on the economy depends upon the difference between the benefits and costs. As mentioned, the depletion allowance, total net operating expenses and lastly the severance and ad valorem expenses did not form part of the analysis.

8.4. Conclusion

The chapter focused on the financial modelling analysis. As mentioned, shale gas, production is an attractive business. There needs to be an accurate means of predicting or forecasting the production volumes to provide to investors to determine whether exploration within a shale play is a lucrative investment. It has been outlined that the EUR, which is obtained from DCA models, plays a major role in the economic or financial analysis of shale gas developments or projects.

During the evaluation process of this study, it was found that, different EUR values are obtained from the various models. The EUR either is over or under-estimated. The ANN and Arps-PLE hybrid decline models, which were found to predict values closest to the actuals, was used to calculate the EUR and compare with other DCA models. The results clearly show the overestimation of the EUR values for the various shale plays, using the Arps, Duong’s and PLE decline models, compared to the ANN and Arps-PLE models. The EUR values obtained from this study were also compared with results from previous studies. The comparison showed an agreement in the results. This would confirm the effectiveness of the models to evaluate the EUR for a shale play.

Evaluating the EUR accurately would then allow for the accurate estimation in the total net revenue generated from a shale play. It is this information, which would attract investors to the development or project. The results from the well economics indicate the lucrative investment shale gas exploration contributes to the economy. However, it has to be highlighted that the total revenue generated would be sensitive to the natural gas price. The higher prices would extend the lifespan of a shale play that would translate to higher generated cash flows for the investor. It is evident that investment in shale gas exploration is precarious, and that one would have to consider the risks associated with the development and their governance, which is excluded from the scope of this research.

Chapter 9

Conclusions and Recommendations

9.1. Introduction

The chapter will summarise the key findings from this study with some recommendations. Initially in this study, an evaluation and sensitivity analysis was conducted on the different decline models. The evaluation process indicated that although there are advantages among the models i.e., the Arps method is simplistic with regard to data fitting, which makes it appealing compared to other decline curve models. However, there are other associated limitations i.e., either there are optimistic or adverse solutions for unconventional reservoirs. During the assessment of the models, the ANN and ARIMA models were also investigated, along with the ANN-ARIMA hybrid model.

The literature has indicated that hybrid models provide a higher degree of accuracy compared to single or individual models (Faruk, 2010). Therefore, due to the strengths and weaknesses associated with the various decline models, trying the hybrid approach was recommended. This would allow for combining the strengths of the individual models while improving their accuracy.

Secondly, this study involved the development of hybrid decline models. The models were chosen according to the sensitivity analysis conducted. Due to the limitation in obtaining shale gas production data, a secondary data collection process was employed. The proposed methodology would assist in obtaining data for large shale plays across the US and provide a comparative basis with this study. Thirdly, the research involved using the data to forecast, determine the accuracy and validate the various models, using the software identified in the research. Lastly, based on the results from the validation process, the model which provided the most accurate predicted values, was used for the economic analysis.

9.2. Conclusions

In this section the five objectives set out for the study will be discussed and reviewed.

Objective 1: Analyse the accuracy and sensitivity of Decline Curve Models

Based on an analysis of previous studies conducted; it was found that the Arps hyperbolic decline, the MHD and Duong's models provided the best fit with production data (Manda and Nkazi, 2020). However, contrary to the reviewed studies it was found that when estimated production data was used in the evaluation process for the purpose of this paper, using the goodness-of-fit technique, the PLE and Duong's decline models aligned the best with the production data, compared to the other models (Manda and Nkazi, 2020).

Studies revealed that the production time significantly affects the EUR, depending on which decline model is used (Manda and Nkazi, 2020). When each model was assessed for accuracy, once again using the goodness-of-fit technique, the results indicated that the SEDM, followed by the LGM, EEDM, PLE, Duong's decline model and, lastly, the hyperbolic decline model align with the production data (Manda and Nkazi, 2020). Based on the goodness-of-fit assessment, the Arps, PLE and Duong's decline models were chosen to evaluate the hybrid model approach, due to the simplicity of the Arps model, while PLE and Duong's model aligned best with the production data.

Objective 2: Develop hybrid decline curve models

During the model development process, the following hybrid decline models were developed for this study:

- Arps-Duong's-PLE Hybrid Decline Model,
- Arps-Duong's Hybrid Decline Model,
- Arps-PLE Hybrid Decline Model and lastly,
- Duong's-PLE Hybrid Decline Model.

Objective 3: Validate the developed model against other DCA models i.e. the three models that have been identified for this study.

The DCA hybrid models, Arps-Duong's-PLE, Arps-Duong's and Duong's-PLE did not give the best outcome as it had been assumed it would, in comparison with the individual DCA models. However, the Arps-PLE hybrid decline model gave the closest predicted results. Both the ARIMA and ANN models gave good, predicted results compared to the DCA models evaluated in this study. However, when both models were combined into the ANN-ARIMA hybrid model, the strengths of both models referenced in the literature did not provide accurate predictive data. The resultant was an overestimation of the production flow rate. This may be attributed to using one input in the ANN model.

Objective 4: Determine which of the decline models accurately forecasts the production decline of shale gas reservoirs

Based on the accuracy and validation results for each model, which ranks the models based on the RMSE value from smallest to highest, the smaller the RMSE value, the better the model. The Arps-PLE hybrid decline model proved the best model i.e., the lowest RMSE value followed by the ANN and ARIMA models.

It is apparent that by combining the Arps and PLE decline models, the limitations in the models are minimised. There is a significant reduction in RMSE values for the individual PLE and Arps models when the models are combined into a hybrid decline model.

The contributing factor that was highlighted earlier is the dominance of the PLE parameters in the model i.e., D_i *and* D_∞ which considers the TFR flows not catered for in the Arps decline model. In addition, the PLE model was specifically developed for SGRs and by combining the models, the number of variables are reduced from four to three. This reduces the number of variables to solve, which was identified as a limitation of the model and hence the degrees of freedom is reduced.

When considering the validation parameters RMSE, MAPE and R^2 the results indicated that, the overall, ANN outperforms all the other methods. The finding concurs with Dhini et al. (2015). They attributed the accuracy due to the non-linearity, however, contrary to their assumption, the ARIMA model proved to be the second most accurate model.

As mentioned by Khashei and Bijari (2010), traditional methods such as the Box–Jenkins model and ARIMA require the assumption that the time series data used in forecasting is linear, and therefore they are not suitable for predicting data that is non-linear.

The reasoning for the ARIMA working well for predicting decline behaviour may be attributed to various flow regimes that can occur in a hydraulically fractured reservoir. The various flow regimes do follow a linear trend. It can be assumed that the weak performance of the other decline curve hybrid models is due to the changes in flow regimes; an effect which is not considered. In addition, one would have to consider extending the data range over a longer period, to determine whether the hybrid models would yield results that are more positive.

Objective 5: Evaluate the economic analysis

Lastly, the economic or financial analysis of this study found that, different EUR values are obtained from the various models. The ANN model, which was found to be the best model in predicting values closest to the actuals, was used to calculate the EUR and compared with other DCA models. The comparison showed agreement in the results. This confirms the effectiveness of the Arps-PLE hybrid decline and ANN models that are used to evaluate the EUR for a shale play.

Evaluating the EUR accurately would then allow for an accurate estimation of the total net revenue generated from a shale play. It is this information, which would attract investors to a development or project. However, it has to be highlighted that the total revenue generated would be sensitive to the natural gas price. The higher prices will extend the lifespan of the shale play that, in turn, would translate to higher generated cash flows for the investor.

In conclusion, the findings have provided a significant contribution to the prediction of shale gas production. The results indicate that the Arps-PLE hybrid decline model is a good model predictor for shale gas production and ultimately the EUR.

9.3. Recommendations

To improve the accuracy of model prediction, the following is recommended for future research:

- The project focus was to establish the accuracy of the hybrid univariant model. Future work will assess the hybrid multi-variant model. Such work will address the problem highlighted in the study, where the ARIMA-ANN model did not provide accurate predictive results which was contrary to the work done by Dhini et al. (2015).
- Researchers should Investigate LSTM as an alternative method to ANN. The model tends to study long-term dependencies and solve the vanishing gradient problems; an issue observed with the ANN model (Tadjer et al., 2021). This would address a factor which contributed to the ANN model not providing good accuracy.
- Researchers should also explore hybrid methods to predict flow regime changes. Kuila et al. (2013) showed that gas flow in SGRs is defined by a combination of mechanisms acting at different scales. According to Huang et al. (2015) gas flow regimes can be classified into four groups, depending on the Knudsen number. Future research should investigate accounting for and incorporating the Knudsen diffusion into the Arps−PLE hybrid decline model in shale gas reservoir modelling.

Table 9.1: Summary of major conclusions and recommendations from the research

	Conclusions	Recommendations
1	The Arps-PLE hybrid decline model gave the closest predicted results.	Assess multivariant hybrid models.
2	The Arps'-PLE hybrid decline model proved to be the best model i.e., lowest RMSE value, followed by the ANN and ARIMA models.	Investigate LSTM an alternative method to ANN.
3	The Arps-PLE hybrid decline model is a good predictor for shale gas production and ultimately the EUR.	Explore hybrid methods to predict flow regime changes, incorporating the Knudsen diffusion into the Arps−PLE hybrid decline model.

References

1. Adekoya, F. (2009). *Production decline analysis of horizontal well in gas shale reservoirs.* West Virginia University.
2. Ali, T. A. & Sheng, J. J. (2015, October). Production decline models: A comparison study. In *SPE Eastern Regional Meeting*. Society of Petroleum Engineers.
3. Arps, J. J. (1945). Analysis of decline curves. *Transactions of the AIME*, *160*(1), 228-247.
4. Ayub, S. & Jafri, Y. Z. (2020). Comparative Study of an ANN-ARIMA Hybrid Model for Predicting Karachi Stock Price. *American Journal of Mathematics and Statistics*, *10*(1), 1–9.
5. Babu, C. N. & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. Applied *Soft Computing*, *23*, 27-38.
6. Bacha, H. & Meyer, W. (1992, June). A neural network architecture for load forecasting. In *IJCNN International Joint Conference on Neural Networks*. IEEE.
7. Bagozzi, R. P. & Yi, Y. (1988). On the evaluation of structural Equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74-94.
8. Boah, E. A., Borsah, A. A. & Brantson, E. T. (2018). Decline Curve Analysis and Production Forecast Studies for Oil Well Performance Prediction: A Case Study of Reservoir X. *International Journal of Engineering Science (IJES)*, *7*, 56-67.
9. Brantson, E. T., Ju, B., Ziggah, Y. Y., Akwensi, P. H., Sun, Y., Wu, D. & Addo, B. J. (2019). Forecasting of horizontal gas well production decline in unconventional reservoirs using productivity, soft computing and swarm intelligence models. *Natural Resources Research*, *28*(3), 717-756.
10. Browning, J., Tinker, S. W., Ikonnikova, S., Gülen, G., Potter, E., Fu, Q. & Roberts, F. (2013). Barnett Shale Well Model-2 (Conclusion): Barnett study

determines full-field reserves, production forecast. *Oil & Gas Journal*, *111*(9), 88-95.

11. Can, B. & Kabir, C. S. (2011, January). Probabilistic performance forecasting for unconventional reservoirs with stretched-exponential model. *In North American Unconventional Gas Conference and Exhibition*. Society of Petroleum Engineers.
12. Chen, Y., Jackson, D. A. & Harvey, H. H. (1992). A comparison of von Bertalanffy and polynomial functions in modelling fish growth data. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*(6), 1228-1235.
13. Chakraborty, K., Mehrotra, K., Mohan, C.K. & Ranka, S. (1992). Forecasting the behaviour of multivariate time series using neural networks. *Neural Networks*, *5*(6), 961-970.
14. Clark, A. J. (2011a). *Decline curve analysis in unconventional resource plays using logistic growth models* (Doctoral dissertation, University of Texas).
15. Clark, A. J., Lake, L. W. & Patzek, T. W. (2011b, January). Production forecasting with logistic growth models. *In SPE annual technical conference and exhibition*. Society of Petroleum Engineers.
16. Contreras, J., Espinola, R., Nogales, F. J. & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, *18*(3), 1014-1020.
17. Cottrell, M., Girard, B., Girard, Y., Mangeas, M. & Muller, C. (1995). Neural modelling for time series: a statistical stepwise method for weight elimination. *IEEE Transactions on Neural Networks*, *6*(6), 1355-1364.
18. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, *2*(4), 303-314.
19. Dayal, A.M. (2017). Environmental concerns of shale gas production. In *Shale Gas* (pp 137-144). Elsevier.

20. De Groot, C. & Würtz, D. (1991). Analysis of univariate time series with connectionist nets: A case study of two classical examples. *Neurocomputing*, *3*(4), 177-192.
21. Dhini, A., Surjandari, I., Riefqi, M. & Puspasari, M. A. (2015). Forecasting analysis of consumer goods demand using neural networks and ARIMA. *International Journal of Technology*, *6*(5), 872-880.
22. Duong, A. N. (2011). Rate-decline analysis for fracture-dominated shale reservoirs. *SPE Reservoir Evaluation & Engineering*, *14*(3), 377-387.
23. Dong, Z., Holditch, S. & McVay, D. (2013). Resource evaluation for shale gas reservoirs. *SPE Economics & Management*, *5*(1), 5-16.
24. Elsaig, M. (2016). *Characterizations of the Marcellus Shale Well petrophysical properties* (Masters dissertation, West Virginia University).
25. Faruk, D. Ö. (2010). A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence*, *23*(4), 586-594.
26. Fattah, J., Ezzine, L., Aman, Z., El Moussami, H. & Lachhab, A. (2018). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management*, *10*, 1-9.
27. Fethney, J. (2010). Statistical and clinical significance and how to use confidence intervals to help interpret both. *Australian Critical Care*, *23*(2), 93-97.
28. Fetkovich, M. J. (1980). Decline curve analysis using type curves. *Journal of Petroleum Technology*, *32*(6), 1-065.
29. Foster, W.R., Collopy, F. & Ungar, L.H., (1992). Neural network forecasting of short, noisy time series. *Computers & Chemical Engineering*, *16*(4), 293-297.
30. Ginzberg, I. & Horn, D. (1991, November). Learnability of times series. In *IEEE International Joint Conference on Neural Networks*. IEEE.

31. Gorr, W.L., Nagin, D. & Szczypula, J. (1994). Comparative study of artificial neural network and statistical models for predicting student grade point averages. *International Journal of Forecasting*, *10*(1), 17-34.
32. Granger, C. W. J. (1989). Combining Forecasts - Twenty Years Later. *Essays in Econometrics: Collected Papers of Clive WJ Granger*, *32*, 411.
33. Grudnitski, G. & Osburn, L. (1993). Forecasting S&P and gold futures prices: An application of neural networks. *Journal of Futures Markets*, *13*(6), 631-643.
34. Guarnone, M., Rossi, F., Negri, E., Grassi, C., Genazzi, D. & Zennaro, R. (2012). An unconventional mindset for shale gas surface facilities. *Journal of Natural Gas Science and Engineering*, *6*, 14-23.
35. Guo, K., Zhang, B., Wachtmeister, H., Aleklett, K. & Höök, M. (2017). Characteristic Production Decline Patterns for Shale Gas Wells in Barnett. *International Journal of Sustainable Future for Human Security*, *5*, 11-20.
36. Harris, S. C. (2013). *A Study of Decline Curve Analysis in the Elm Coulee Field* (Masters dissertation, Texas A & M University).
37. Hill, T., O'Connor, M. & Remus, W. (1996). Neural network models for time series forecasts. *Management Science*, *42*(7), 1082-1092.
38. Howarth, R. W., Santoro, R. & Ingraffea, A. (2011). Methane and the greenhouse-gas footprint of natural gas from shale formations. *Climatic Change*, *106*(4), 679.
39. Hubbert, M.K. & Willis, D.G. (1957). Mechanics of hydraulic fracturing. *Transactions of the AIME*, *210*(1), 153-168.
40. Ilk, D., Rushing, J. A., Perego, A. D. & Blasingame, T. A. (2008, January). Exponential vs. hyperbolic decline in tight gas sands: understanding the origin and implications for reserve estimates using Arps decline curves. In *SPE annual technical conference and exhibition*. Society of Petroleum Engineers.

41. Johnson, N. L., Currie, S. M., Ilk, D. & Blasingame, T. A. (2009, January). A simple methodology for direct estimation of gas-in-place and reserves using rate-time data. In *SPE Rocky Mountain Petroleum Technology Conference*. Society of Petroleum Engineers.
42. Joshi, K. J. (2012). *Comparison of Various Deterministic Forecasting Techniques in Shale Gas Reservoirs with Emphasis on the Duong Method* (Doctoral dissertation, Texas A & M University).
43. Kaastra, I. & Boyd, M. (1996). Designing a neural network for forecasting financial. *Neurocomputing*, *10*, 215-236.
44. Kanfar, M.S. (2013). *Comparison of Empirical Decline Curve Analysis for Shale Wells* (Doctoral dissertation, Texas A & M University).
45. Kanfar, M. & Wattenbarger, R. (2012, January). Comparison of empirical decline curve methods for shale wells. In *SPE Canadian Unconventional Resources Conference*. Society of Petroleum Engineers.
46. Kang, S.Y. (1991). *An investigation of the use of feedforward neural networks for forecasting* (Doctoral dissertation, Kent State University).
47. Kargbo, D. M., Wilhelm, R. G. & Campbell, D. J. (2010). Natural gas plays in the Marcellus Shale Well: Challenges and potential opportunities. 5679-5684.
48. Kendon, P. (2019). Bottom Hole Treatment Pressure: What It Means to Your Fracking Project. *Trenchlesspedia*
49. Kenomore, M., Hassan, M., Malakooti, R., Dhakal, H. & Shah, A. (2018). Shale gas production decline trend over time in the Barnett Shale Well. *Journal of Petroleum Science and Engineering*, *165*, 691-710.
50. Kinnaman, T. C. (2011). The economic impact of shale gas extraction: A review of existing studies. *Ecological Economics*, *70*(7), 1243-1249.
51. Kisslinger, C. (1993). The stretched exponential function as an alternative model for aftershock decay rate. *Journal of Geophysical Research: Solid Earth*, *98*(B2), 1913-1921.

52. Khashei, M. & Bijari, M. (2011). A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing*, *11*(2), 2664-2675.
53. Knudsen, B. R., Foss, B., Whitson, C. H. & Conn, A. R. (2012). Target-rate tracking for shale-gas multi-well pads by scheduled shut-ins. *IFAC Proceedings Volumes*, *45*(15), 107-113.
54. Kohzadi, N., Boyd, M.S., Kermanshahi, B. & Kaastra, I. (1996). A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing*, *10*(2), 169-181.
55. Kothari, C. R. (2004). *Research methodology: Methods and techniques*. New Age International.
56. Kuan, C.M. & Liu, T. (1995). Forecasting exchange rates using feedforward and recurrent neural networks. *Journal of applied econometrics*, *10*(4), 347-364.
57. Lachtermacher, G. & Fuller, J.D. (1995). Back propagation in time-series forecasting. *Journal of forecasting*, *14*(4), 381-393.
58. Lake, L. W., Martin, J., Ramsey, J. D. & Titman, S. (2013). A primer on the economics of shale gas production just how cheap is shale gas?. *Journal of Applied Corporate Finance*, *25*(4), 87-96.
59. Leblanc, D. & Okouma, V. (2018). New rate decline EUR model for unconventional reservoirs. *World Oil*, *239*(3), 55-58.
60. Lee, K. S. & Kim, T. H. (2016). *Integrative understanding of shale gas reservoirs* (p. 123). Springer International Publishing.
61. Li, P., Hao, M., Hu, J., Ru, Z. & Li, Z. (2018). A new production decline model for horizontal wells in low-permeability reservoirs. *Journal of Petroleum Science and Engineering*, *171*, 340-352.
62. Loh, H. P. & Loh, N. (2016). Hydraulic fracturing and shale gas: Environmental and health impacts. In *Advances in Water Resources Management* (pp. 293-337). Springer, Cham.

63. Nam, K. & Schaefer, T. (1995). Forecasting international airline passenger traffic using neural networks. *The Logistics and Transportation Review*, *31*(3), 239-252.

64. Nelson, M., Hill, T., Remus, B. & O'Connor, M. (1994, January). Can neural networks applied to time series forecasting learn seasonal patterns: an empirical investigation. *In the Twenty-Seventh Hawaii International Conference on System Sciences*.

65. Nelson, P. H. (2009). Pore-throat sizes in sandstones, tight sandstones, and shales. *AAPG Bulletin*, *93*(3), 329-340.

66. Makinde, F. A., Orodu, O. D., Ladipo, A. O. & Anawe, P. A. L. (2012). Cumulative Production Forecast of an Oil Well Using Simplified Hyperbolic-Exponential Decline Models. *Global Journal of Research in Engineering*, *12*(2-B).

67. Makinde, I. & Lee, W. J. (2017). Forecasting production of liquid rich shale (LRS) reservoirs using simple models. *Journal of Petroleum Science and Engineering*, *157*, 461-481.

68. Manda, P. & Nkazi, D. B. (2020). The Evaluation and Sensitivity of Decline Curve Modelling. *Energies*, *13*(11), 2765.

69. McCarthy, N. (2017). The American Shale Revolution, Available at: www.statista.com/chart/11830/the-american-shale-revolution/, (Accessed on 14 March 2019)

70. McNeil, R., Jeje, O. & Renaud, A. (2009, January). Application of the power law loss-ratio method of decline analysis. In *Canadian International Petroleum Conference*. Petroleum Society of Canada.

71. Mittag-Leffler, G. M. (1903). Sur la nouvelle function Eα(x), *Comptes rendus de l'Académie des Sciences, Paris*, *137*, 554–558.

72. Moridis, G. J., Reagan, M. T., Kuzma, H. A., Blasingame, T. A., Huang, Y. W., Santos, R. & Bhattacharya, S. (2013). SeTES: A self-teaching expert system

for the analysis, design, and prediction of gas production from unconventional gas resources. *Computers & Geosciences*, *58*, 100-115.

73. Nwaobi, U. & Anandarajah, G. (2018). A Critical Review of Shale Gas Production Analysis and Forecast Methods. *Saudi Journal of Engineering and Technology (SJEAT)*.

74. Paryani, M., Awoleke, O. O., Ahmadi, M., Hanks, C. & Barry, R. (2017). Approximate Bayesian Computation for Probabilistic Decline-Curve Analysis in Unconventional Reservoirs. *SPE Reservoir Evaluation & Engineering*, *20*(2), 478-485.

75. Paryani M., Ahmadi M., Awoleke O. & Hanks C. (2018). Decline Curve Analysis: A Comparative Study of Proposed Models Using Improved Residual Functions. *Journal of Petroleum and Environmental Biotechnology*, *9*(1), 1-8.

76. Patzek, T. W., Male, F. & Marder, M. (2013). Gas production in the Barnett Shale Well obeys a simple scaling theory. *Proceedings of the National Academy of Sciences*, *110*(49), 19731-19736.

77. Penner, S. S., Alpert, S. B. & Bendanillo, V. (2013). *New Sources of Oil and Gas: Gases from Coal; Liquid Fuels from Coal, Shale, Tar Sands, and Heavy Oil Sources*. Elsevier.

78. Perrone, M. P., & Cooper, L. N. (1992). When networks disagree: Ensemble methods for hybrid neural networks. *Brown University Providence RI Institute for Brain and Neural Systems.*

79. Qu, Z. (Ed.). (2018). *Proceedings of the International Field Exploration and Development Conference*. Springer.

80. Raymond, Y. C. (1997). An application of the ARIMA model to real-estate prices in Hong Kong. *Journal of Property Finance*, *8*(2), 152-163.

81. Reilly, D.L. & Cooper, L.N. (1995). An overview of neural networks: early models to real world systems. *How We Learn; How We Remember: Toward an*

Understanding of Brain and Neural Systems: Selected Papers of Leon N Cooper, 300-321.

82. Robertson, S. (1988). Generalised Hyperbolic Equation. Society of Petroleum Engineers, Richardson, Texas, USA, 1988.
83. Sagheer, A. & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, *323*, 203-213.
84. Schöneburg, E. (1990). Stock price prediction using neural networks: A project report. *Neurocomputing*, *2*(1), 17-27.
85. Seshadri, J. N. & Mattar, L. (2010, January). Comparison of power law and modified hyperbolic decline methods. In *Canadian Unconventional Resources and International Petroleum Conference*. Society of Petroleum Engineers.
86. Shah, S. (2013). *Development of New Decline Model for Shale Oil Reserves* (Doctoral dissertation, University of Houston).
87. Shamsuddin, S. M., Sallehuddin, R. & Yusof, N. M. (2008). Artificial neural network time series modelling for revenue forecasting. *Chiang Mai Journal of Science*, *35*(3), 411-426.
88. Sharda, R. & Patil, R.B. (1992). Connectionist approach to time series prediction: an empirical test. *Journal of Intelligent Manufacturing*, *3*(5), 317-323.
89. Shukla, M. & Jharkharia, S. (2011). ARIMA models to forecast demand in fresh supply chains. *International Journal of Operational Research*, *11*(1), 1-18.
90. Soomro, H. T., Memon, A & Lashari, Z. A. (2017). Environmental Impacts of Shale Gas Exploitation. *Australian Journal of Engineering and Technology Research*, *2*(1), 73-80
91. Srinivasan, D., Liew, A.C. & Chang, C.S. (1994). Forecasting daily load curves using a hybrid fuzzy-neural approach. *IEE Proceedings-Generation, Transmission and Distribution*, *141*(6), 561-567.

92. Steyl, G., Van Tonder, G. J. & Chevallier, L. P. (2012). State of the Art: Fracking for Shale Gas Exploration in South Africa and the Impact on Water Resources: Report to the Water Research Commission. Water Research Commission.
93. Swindell, G.S. (2018). Estimated Ultimate Recovery (EUR) Study of 5,000 Marcellus Shale Wells in Pennsylvania, http://gswindell.com.
94. Tadjer, A., Hong, A. & Bratvold, R.B. (2021). Machine learning based decline curve analysis for short-term oil production forecast. *Energy Exploration & Exploitation*, 1-23.
95. Tan, L., Zuo, L. & Wang, B. (2018). Methods of decline curve analysis for shale gas reservoirs. *Energies*, *11*(3), 552.
96. Tang, Z., De Almeida, C. & Fishwick, P.A. (1991). Time series forecasting using neural networks vs. Box-Jenkins methodology. *Simulation*, *57*(5), 303-310.
97. Tang, Z. & Fishwick, P.A. (1993). Feedforward neural nets as models for time series forecasting. *ORSA Journal on Computing*, *5*(4), 374-385.
98. Taskaya-Temizel, T. & Ahmad, K. (2005, July). Are ARIMA neural network hybrids better than single models?. In *IEEE International Joint Conference on Neural Networks*. IEEE.
99. Theodori, G. L. (2013). Perception of the natural gas industry and engagement in individual civic actions. *Journal of Rural Social Sciences*, *28*(2), 5.
100. Tsoularis, A. & Wallace, J. (2002). Analysis of logistic growth models. *Mathematical Biosciences*, *179*(1), 21-55.
101. Wait, R. & Rossouw, R. (2014). A comparative assessment of the economic benefits from shale gas extraction in the Karoo, South Africa. *Southern African Business Review*, *18*(2), 1-34.
102. Valkó, P. P. (2009, January). Assigning value to stimulation in the Barnett Shale Well: a simultaneous analysis of 7000 plus production hystories and well completion records. *In SPE hydraulic fracturing technology conference.* Society of Petroleum Engineers.

103. Valkó, P. P. & Lee, W. J. (2010, January). A better way to forecast production from unconventional gas wells. *In SPE annual technical conference and exhibition*. Society of Petroleum Engineers.

104. Vanorsdale, C. R. (2013, September). Production decline analysis lessons from classic shale gas wells. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

105. Vishwakarma, K.P. (1994, June). A neural network to predict multiple economic time series. *In IEEE International Conference on Neural Networks*. IEEE.

106. Wachtmeister, H., Lund, L., Aleklett, K. & Höök, M. (2017). Production decline curves of tight oil wells in eagle ford shale. *Natural Resources Research*, *26*(3), 365-377.

107. Wait, R. & Rossouw, R. (2014). A comparative assessment of the economic benefits from shale gas extraction in the Karoo, South Africa. *Southern African Business Review*, *18*(2), 1-34.

108. Wang, H. (2017). What factors control shale gas production and production decline trend in fractured systems: a comprehensive analysis and investigation arXiv preprint arXiv:1710.11464.

109. Wang, L., Tian, Y., Yu, X., Wang, C., Yao, B., Wang, S. & Cui, J. (2017). Advances in improved/enhanced oil recovery technologies for tight and shale reservoirs. *Fuel*, *210*, 425-445.

110. Wang, K., Li, H., Wang, J., Jiang, B., Bu, C., Zhang, Q. & Luo, W. (2017). Predicting production and estimated ultimate recoveries for shale gas wells: A new methodology approach. *Applied Energy*, *206*, 1416-1431.

111. Weigend, A. (1992). Predicting sunspots and exchange rates with connectionist networks. *Nonlinear Model Forecasting Journal*. 395-432.

112. Weijermars, R. (2013). Economic appraisal of shale gas plays in Continental Europe. *Applied Energy*, *106*, 100-115.

113. Weijermars, R., van Harmelen, A., Zuo, L., Nascentes, I.A. & Yu, W. (2017a), July. High-resolution visualization of flow interference between frac clusters (part 1): model verification and basic cases. *In SPE/AAPG/SEG Unconventional Resources Technology Conference*. OnePetro.

114. Weijermars, R., van Harmelen, A. & Zuo, L. (2017b), September. Flow interference between fracture clusters (part 2): field example from the Midland basin (Wolfcamp Formation, Spraberry Trend Field) with implications for hydraulic fracture design. In *Unconventional Resources Technology Conference.* Society of Exploration Geophysicists, American Association of Petroleum Geologists, Society of Petroleum Engineers.

115. Weijermars, R. & Nandlal, K. (2020). Pre-Drilling Production Forecasting of Parent and Child Wells Using a 2-Segment Decline Curve Analysis (DCA) Method Based on an Analytical Flow-Cell Model Scaled by a Single Type Well. *Energies*, *13*(6), 1525.

116. Yuan, J., Luo, D. & Feng, L. (2015). A review of the technical and economic evaluation techniques for shale gas development. *Applied Energy*, *148*, 49-65.

117. Yuan, Y., Qi, Z., Chen, Z., Yan, W. & Zhao, Z. (2020). Production decline analysis of shale gas based on a probability density distribution function. *Journal of Geophysics and Engineering*, *17*(2), pp.365-376.

118. Yuhu, B., Guihua, C., Bingxiang, X., Ruyong, F. & Ling, C. (2016). Comparison of typical curve models for shale gas production decline prediction. *China Petroleum Exploration*, *21*(5), 96-102.

119. Zhang, X. (1994). Time series analysis and prediction by neural networks. *Optimization Methods and Software*, *4*(2), 151-170.

120. Zhang, X. & Numao, M. (1997). MPL-Core: An Efficient Multiple Predicate Learner Based on Fast Failure Mechanism. 人工知能, *12*(4), 582-590.

121. Zhang, G., Patuwo, B. E. & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of forecasting*, *14*(1), 35-62.

122. Zhang, G. P., Patuwo, B. E. & Hu, M. Y. (2001). A simulation study of artificial neural networks for nonlinear time-series forecasting. *Computers & Operations Research*, *28*(4), 381-396.

123. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159-175.

124. Zhang, H., Rietz, D., Cagle, A., Cocco, M. & Lee, J. (2016a). Extended exponential decline curve analysis. *Journal of Natural Gas Science and Engineering*, *36*, 402-413.

125. Zhang, H. E., Nelson, E., Olds, D., Rietz, D. & Lee, W. J. (2016b, September). Effective Applications of Extended Exponential Decline Curve Analysis to both Conventional and Unconventional Reservoirs. *In SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

126. Zhang, X., Wang, X., Hou, X. & Xu, W. (2017). Rate decline analysis of vertically fractured wells in shale gas reservoirs. *Energies*, *10*(10), 1602.

127. Zhou, L. & Selim, H. M. (2003). Application of the fractional advection-dispersion Equation in porous media. *Soil Science Society of America Journal*, *67*(4), 1079-1084.

128. Zuo, L., Yu, W. & Wu, K. (2016). A fractional decline curve analysis model for shale gas reservoirs. *International Journal of Coal Geology*, *163*, 140-148.

Appendices

Appendix A: Raw Data

A.1. Secondary Production Data

Table A1: Estimated raw data and calculated values for the Canon Shale Well within Eagle Ford shale play (Paryani et al., 2018)

Production date (days)	Gas rate, Mscf/d	log qt	log t	qt/qi	qt/t	(1/b)/qt	ln qt/qm
260	20	1.30	2.41	0.007	0.08	0.131	-3.538
240	50	1.70	2.38	0.017	0.21	0.052	-2.622
200	70	1.85	2.30	0.024	0.35	0.037	-2.285
160	90	1.95	2.20	0.031	0.56	0.029	-2.034
140	100	2.00	2.15	0.035	0.71	0.026	-1.929
110	150	2.18	2.04	0.052	1.36	0.017	-1.523
108	200	2.30	2.03	0.069	1.85	0.013	-1.235
105	250	2.40	2.02	0.086	2.38	0.010	-1.012
101	260	2.41	2.00	0.090	2.57	0.010	-0.9731
100	270	2.43	2.00	0.093	2.70	0.010	-0.9358
90	300	2.48	1.95	0.104	3.33	0.009	-0.830
80	370	2.57	1.90	0.128	4.63	0.007	-0.620
70	400	2.60	1.85	0.138	5.71	0.007	-0.542
60	430	2.63	1.78	0.148	7.17	0.006	-0.470
55	450	2.65	1.74	0.155	8.18	0.006	-0.425
53	470	2.67	1.72	0.162	8.87	0.006	-0.381
52	480	2.68	1.72	0.166	9.23	0.005	-0.3603
51	500	2.70	1.71	0.173	9.80	0.005	-0.319
50	600	2.78	1.70	0.207	12.00	0.004	-0.137
49	620	2.79	1.69	0.214	12.65	0.004	-0.104
45	650	2.81	1.65	0.224	14.44	0.004	-0.0568
47	680	2.83	1.67	0.235	14.47	0.004	-0.0117
42	700	2.85	1.62	0.242	16.67	0.004	0.0173
40	720	2.86	1.60	0.248	18.00	0.004	0.0455
36	750	2.88	1.56	0.259	20.83	0.003	0.0863

Table A2: Estimated raw data and calculated values for the Marcellus Shale Well (Adekoya, 2009)

Production date (days)	Gas rate, Mscf/d	log qt	log t	qt/qi	qt/t	(1/b)/qt	ln qt/qm
1000	2500	3.40	3.00	0.436	2.50	0.0004	0.420
1100	2400	3.38	3.04	0.418	2.18	0.0004	0.379
1200	2300	3.36	3.08	0.401	1.92	0.0004	0.336
1500	2000	3.30	3.18	0.349	1.33	0.0005	0.197
1900	1800	3.26	3.28	0.314	0.95	0.0006	0.091
2200	1600	3.20	3.34	0.279	0.73	0.0006	-0.027
2400	1500	3.18	3.38	0.262	0.63	0.0007	-0.091
2600	1400	3.15	3.41	0.244	0.54	0.0007	-0.160
2800	1300	3.11	3.45	0.227	0.46	0.0008	-0.234
3000	1200	3.08	3.48	0.209	0.40	0.0008	-0.314
3200	1100	3.04	3.51	0.192	0.34	0.0009	-0.401
3400	1000	3.00	3.53	0.174	0.29	0.0010	-0.497
3600	900	2.95	3.56	0.157	0.25	0.0011	-0.602
3800	800	2.90	3.58	0.139	0.21	0.0013	-0.720
4000	700	2.85	3.60	0.122	0.18	0.0014	-0.853
4200	600	2.78	3.62	0.105	0.14	0.0017	-1.007
4400	500	2.70	3.64	0.087	0.11	0.0020	-1.190
4600	400	2.60	3.66	0.070	0.09	0.0025	-1.413
4800	300	2.48	3.68	0.052	0.06	0.0033	-1.700
5000	250	2.40	3.70	0.044	0.05	0.0040	-1.883
6000	150	2.18	3.78	0.026	0.03	0.0067	-2.394
7000	100	2.00	3.85	0.017	0.01	0.0100	-2.799
7500	50	1.70	3.88	0.009	0.01	0.0200	-3.492
10000	40	1.60	4.00	0.007	0.00	0.0250	-3.715
12500	30	1.48	4.10	0.005	0.00	0.0333	-4.003
15000	20	1.30	4.18	0.003	0.00	0.0500	-4.409
17500	10	1.00	4.24	0.002	0.00	0.1000	-5.102

Table A3: Estimated raw data and calculated values for the Barnett Shale Well (Tan et al., 2018)

Production date (days)	Gas rate, Mscf/d	log qt	log t	qt/qi	qt/t	(1/b)/qt	ln qt/qm
5	9000	3.95	0.70	1.060	1800.00	0.000090	0.371
8	8000	3.90	0.90	0.942	1000.00	0.000101	0.254
9.7	8500	3.93	0.99	1.001	876.29	0.000095	0.314
9.9	8200	3.91	1.00	0.966	828.28	0.000098	0.278
20	7500	3.88	1.30	0.883	375.00	0.000107	0.189
50	6500	3.81	1.70	0.766	130.00	0.000124	0.046
60	6000	3.78	1.78	0.707	100.00	0.000134	-0.034
80	4000	3.60	1.90	0.471	50.00	0.000202	-0.440
90	3800	3.58	1.95	0.448	42.22	0.000212	-0.491
100	3500	3.54	2.00	0.412	35.00	0.000230	-0.573
200	3000	3.48	2.30	0.353	15.00	0.000269	-0.727
350	2500	3.40	2.54	0.294	7.14	0.000322	-0.910
450	2000	3.30	2.65	0.236	4.44	0.000403	-1.133
600	1500	3.18	2.78	0.177	2.50	0.000537	-1.420
700	1000	3.00	2.85	0.118	1.43	0.000806	-1.826
800	900	2.95	2.90	0.106	1.13	0.000896	-1.931
900	800	2.90	2.95	0.094	0.89	0.001008	-2.049
1000	790	2.90	3.00	0.093	0.79	0.001020	-2.062
1100	780	2.89	3.04	0.092	0.71	0.001033	-2.074
1200	760	2.88	3.08	0.090	0.63	0.001061	-2.100
1300	740	2.87	3.11	0.087	0.57	0.001089	-2.127

Table A4: Estimated raw data and calculated values for Well #A (Brantson et al., 2019)

Production date (days)	Gas rate, Mscf/d	log qt	log t	qt/qi	qt/t	(1/b)/qt	ln qt/qm
5	6500	3.813	0.699	0.881	1300.000	5.09E-05	0.121
7	6400	3.806	0.845	0.867	914.286	5.17E-05	0.105
9	6300	3.799	0.954	0.853	700.000	5.25E-05	0.090
11	6100	3.785	1.041	0.826	554.545	5.43E-05	0.057
20	6500	3.813	1.301	0.881	325.000	5.09E-05	0.121
25	6300	3.799	1.398	0.853	252.000	5.25E-05	0.090
27	6100	3.785	1.431	0.826	225.926	5.43E-05	0.057
28	6100	3.785	1.447	0.826	217.857	5.43E-05	0.057
35	4800	3.681	1.544	0.650	137.143	6.9E-05	-0.182
37	4750	3.677	1.568	0.643	128.378	6.97E-05	-0.193
50	4000	3.602	1.699	0.542	80.000	8.28E-05	-0.365
52	3800	3.580	1.716	0.515	73.077	8.71E-05	-0.416
55	4300	3.633	1.740	0.582	78.182	7.7E-05	-0.292
57	4350	3.638	1.756	0.589	76.316	7.61E-05	-0.281
59	4320	3.635	1.771	0.585	73.220	7.66E-05	-0.288
80	4310	3.634	1.903	0.584	53.875	7.68E-05	-0.290
85	4305	3.634	1.929	0.583	50.647	7.69E-05	-0.291
90	4200	3.623	1.954	0.569	46.667	7.88E-05	-0.316
100	4100	3.613	2.000	0.555	41.000	8.07E-05	-0.340
120	4000	3.602	2.079	0.542	33.333	8.28E-05	-0.365
140	3500	3.544	2.146	0.474	25.000	9.46E-05	-0.498
150	3700	3.568	2.176	0.501	24.667	8.95E-05	-0.443
170	3600	3.556	2.230	0.488	21.176	9.19E-05	-0.470
190	3500	3.544	2.279	0.474	18.421	9.46E-05	-0.498
240	4100	3.613	2.380	0.555	17.083	8.07E-05	-0.340

A.2. KAPPA-Citrine

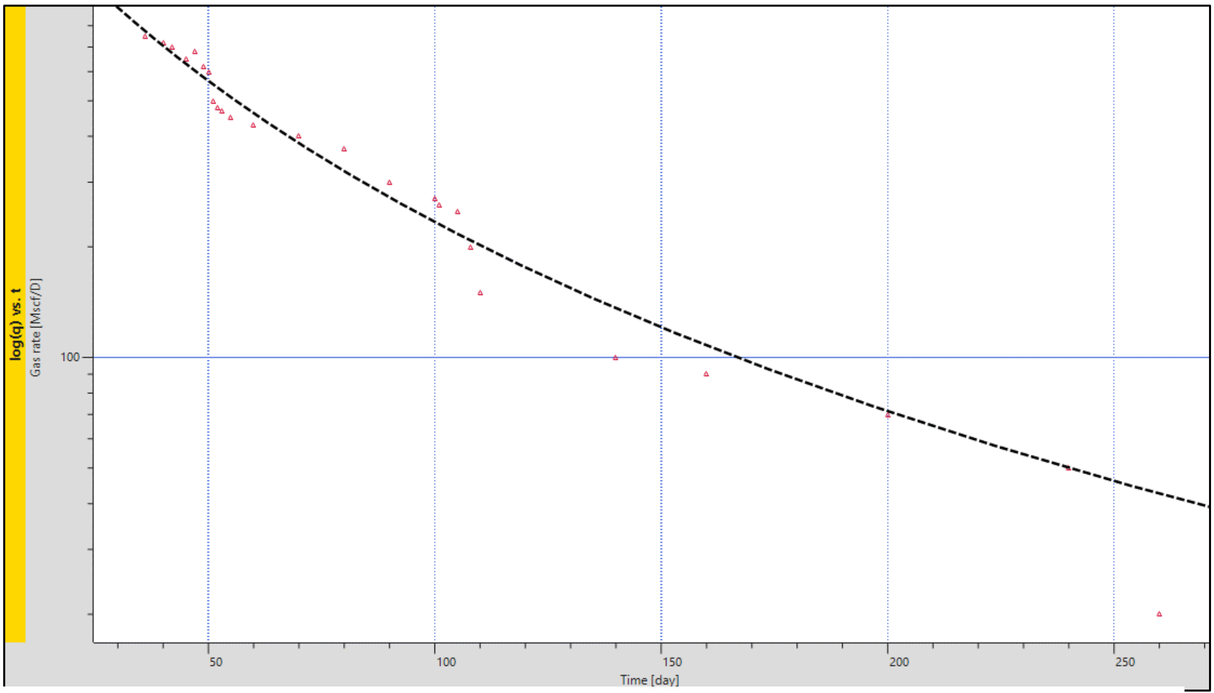


Figure A1: Arps model generated graph from KAPPA-Citrine for the Canon Shale Well (Paryani et al., 2018)

Table A5: Arps model parameters generated by KAPPA-Citrine for the Canon Shale

Parameters	Hyperbolic [2021/06/07]	
Tmin	0.00000	day
Tmax	11210.0	day
Initial rate	2182.90	Mscf/D
Decline parameter	3.53383	%/day
Initial tangent effective decline rate	0.273972	%/day
Initial secant effective decline rate	0.271356	%/day
b exponent	0.383141	
Abandonment rate	0.00435134	Mscf/D
EUR	0.100108	bscf
Relative std. error on rates	11.6231	%
Relative std. error on cumulative	32.1926	%
Remaining Recoverable	0.0336181	bscf

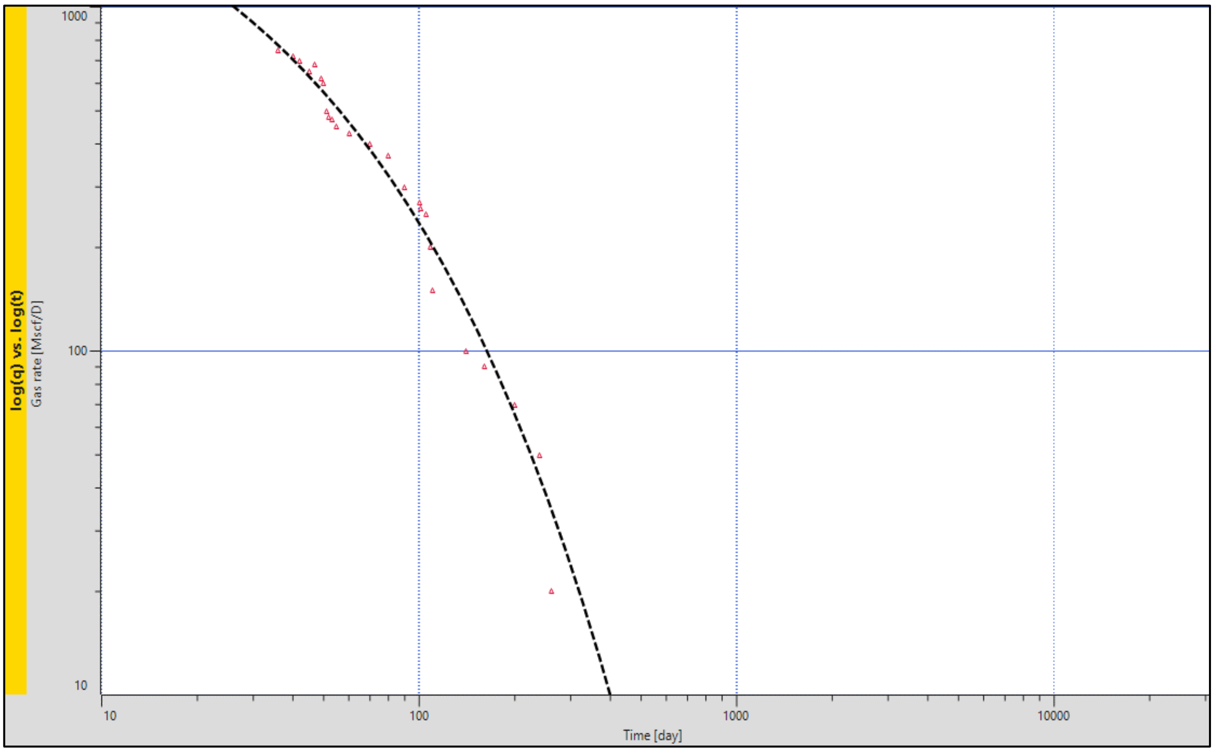


Figure A2: PLE model graph generated by KAPPA-Citrine for the Canon Shale Well (Paryani et al., 2018)

Table A6: PLE model parameters generated by KAPPA-Citrine for the Canon Shale Well

Parameters	Power-law exponential [2021/06/07]	
Tmin	0.00000	day
Tmax	11210.0	day
Initial rate	4018.02	Mscf/D
Decline parameter	6.81226E-4	1/day
Transition time	14.2974	day
n exponent	0.52362	
Abandonment rate	1.13140E-14	Mscf/D
EUR	0.10115	bscf
Relative std. error on rates	11.8283	%
Relative std. error on cumulative	38.0453	%
Remaining Recoverable	0.0346599	bscf

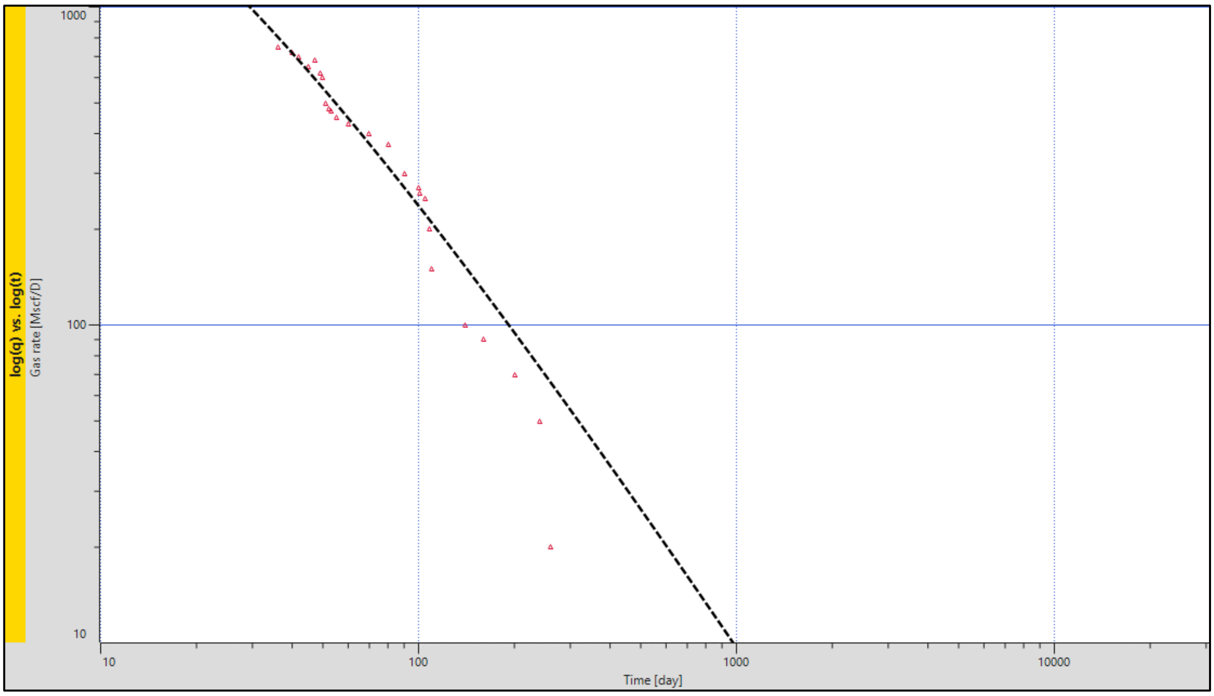


Figure A3: Duong's model graph generated by KAPPA-Citrine for the Canon Shale Well (Paryani et al., 2018)

Table A7: Duong's model parameters generated by KAPPA-Citrine for the Canon Shale Well

Parameters	Duong [2021/06/07]	
Tmin	0.00000	day
Tmax	11210.0	day
Initial rate	2492.89	Mscf/D
M exponent	1.53143	
A exponent	2.74193	
Abandonment rate	0.262992	Mscf/D
EUR	0.151716	bscf
Relative std. error on rates	15.4275	%
Relative std. error on cumulative	51.3421	%
Remaining Recoverable	0.0852256	bscf

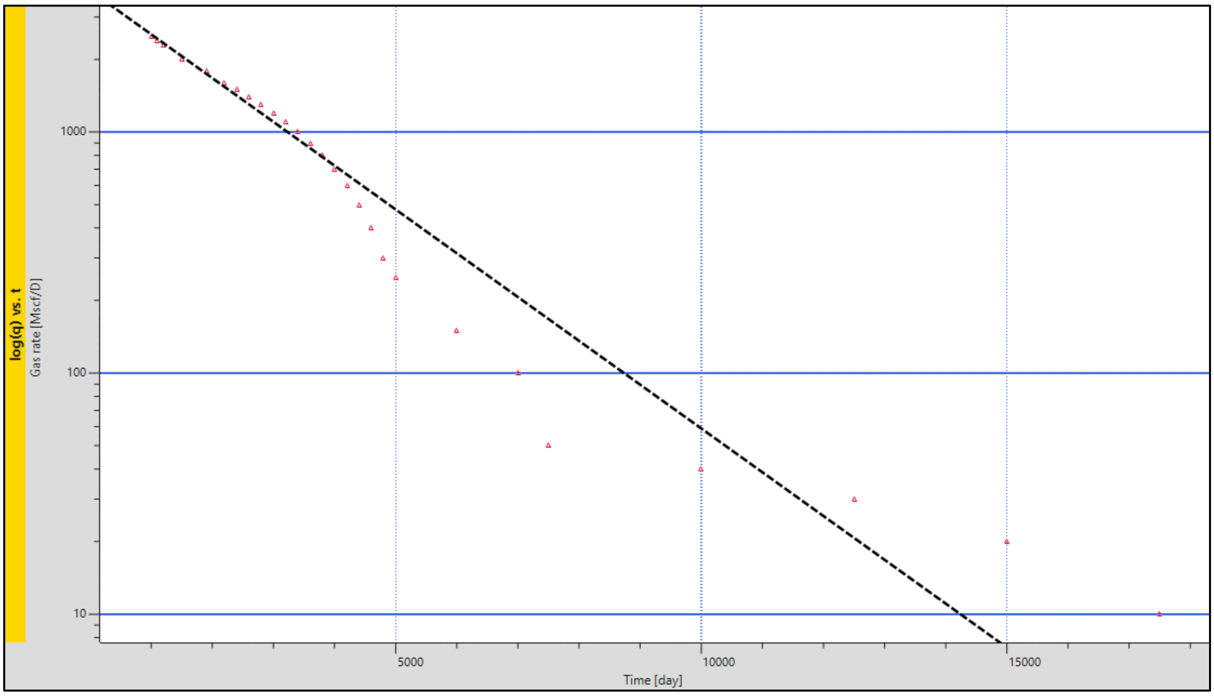


Figure A4: Arps model generated graph from KAPPA-Citrine for the Marcellus Shale Well (Adekoya, 2009)

Table A8: Arp's model parameters generated by KAPPA-Citrine for the Marcellus Shale Well

Parameters	Hyperbolic [2068/08/19]	
Tmin	0.00000	day
Tmax	28450.0	day
Initial rate	3863.51	Mscf/D
Decline parameter	0.0418307	%/day
Initial tangent effective decline rate	0.0387938	%/day
Initial secant effective decline rate	0.0387938	%/day
b exponent	0.00000	
Abandonment rate	0.0262131	Mscf/D
EUR	9.23601	bscf
Relative std. error on rates	34.6712	%
Relative std. error on cumulative	13.8210	%
Remaining Recoverable	1.55102	bscf

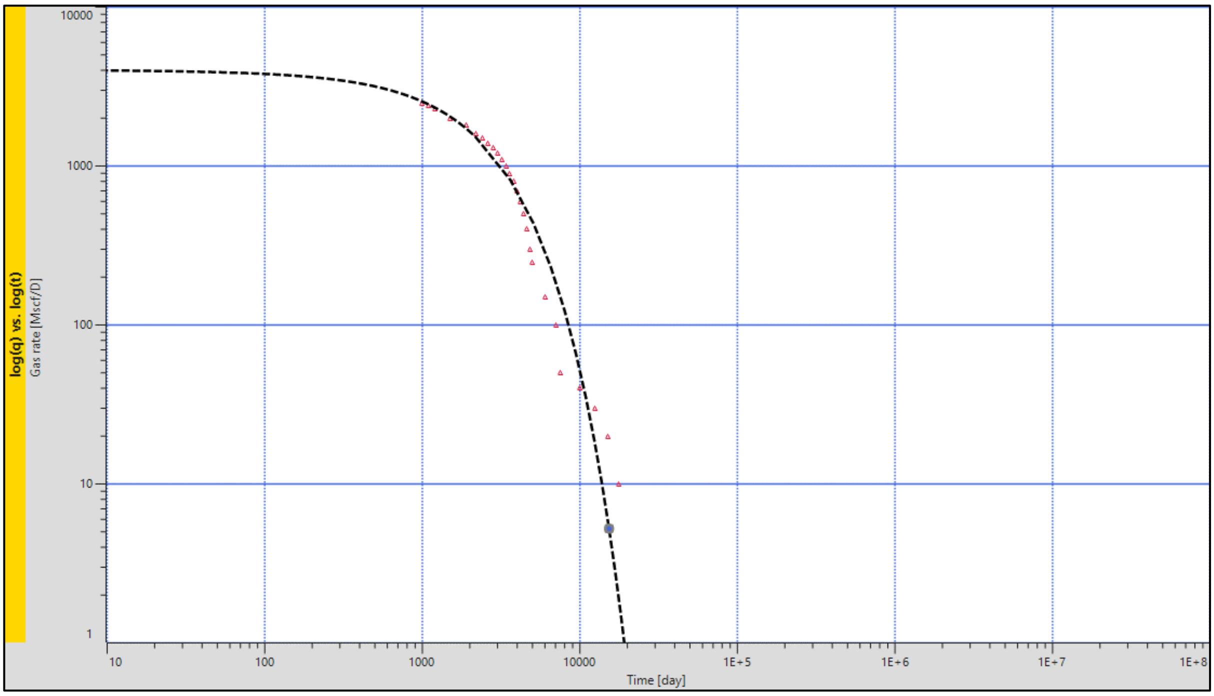


Figure A5: PLE model generated graph from KAPPA-Citrine for the Marcellus Shale Well (Adekoya, 2009)

Table A9: PLE model parameters generated by KAPPA-Citrine for the Marcellus Shale Well

Parameters	Power-law exponential [2068/08/19]	
Tmin	0.00000	day
Tmax	28450.0	day
Initial rate	10812.0	Mscf/D
Decline parameter	4.31451E-4	1/day
Transition time	73.0387	day
n exponent	0.005	
Abandonment rate	0.0180135	Mscf/D
EUR	9.08945	bscf
Relative std. error on rates	41.0778	%
Relative std. error on cumulative	13.6774	%
Remaining Recoverable	1.40445	bscf

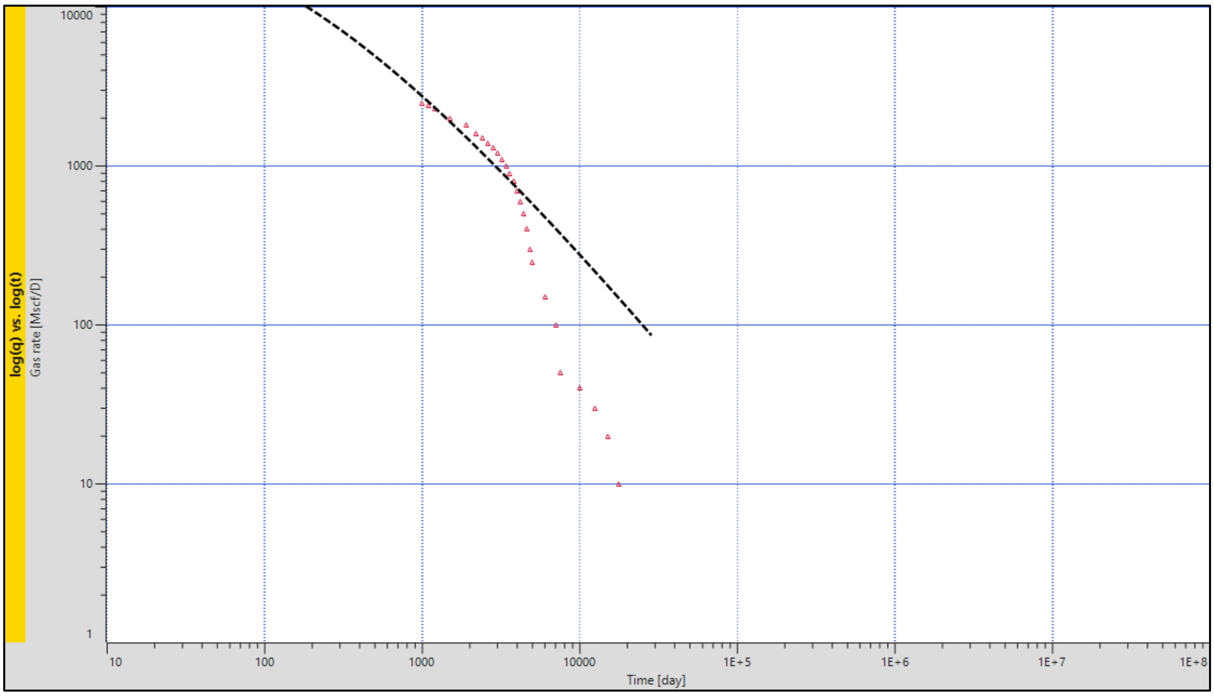


Figure A6: Duong's model generated graph from KAPPA-Citrine for the Marcellus Shale Well (Adekoya, 2009)

Table A10: PLE model parameters generated by KAPPA-Citrine for the Marcellus Shale Well

Parameters	Duong [2068/08/19]	
Tmin	0.00000	day
Tmax	28450.0	day
Initial rate	2529.05	Mscf/D
M exponent	1.29182	
A exponent	3.03243	
Abandonment rate	86.2456	Mscf/D
EUR	16.1366	bscf
Relative std. error on rates	36.2672	%
Relative std. error on cumulative	44.2065	%
Remaining Recoverable	8.45163	bscf

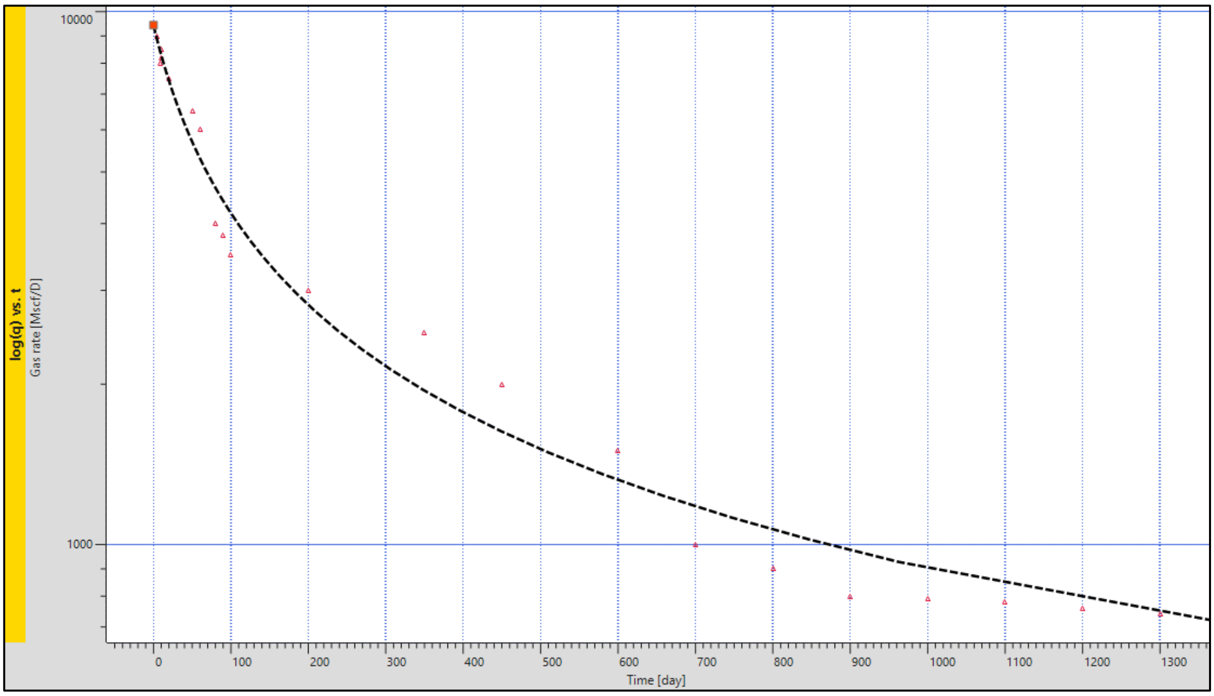


Figure A7: Arps model generated graph from KAPPA-Citrine for the Barnett Shale Well (Tan et al., 2018)

Table A11: Arps model parameters generated by KAPPA-Citrine for the Marcellus Shale Well

Parameters	Hyperbolic [2024/04/12]	
Tmin	0.00000	day
Tmax	12250.0	day
Initial rate	9432.99	Mscf/D
Decline parameter	1.40187	%/day
Initial tangent effective decline rate	0.27233	%/day
Initial secant effective decline rate	0.219123	%/day
b exponent	1.23950	
Abandonment rate	124.384	Mscf/D
EUR	5.11314	bscf
Relative std. error on rates	10.8184	%
Relative std. error on cumulative	3.54923	%
Remaining Recoverable	2.86731	bscf

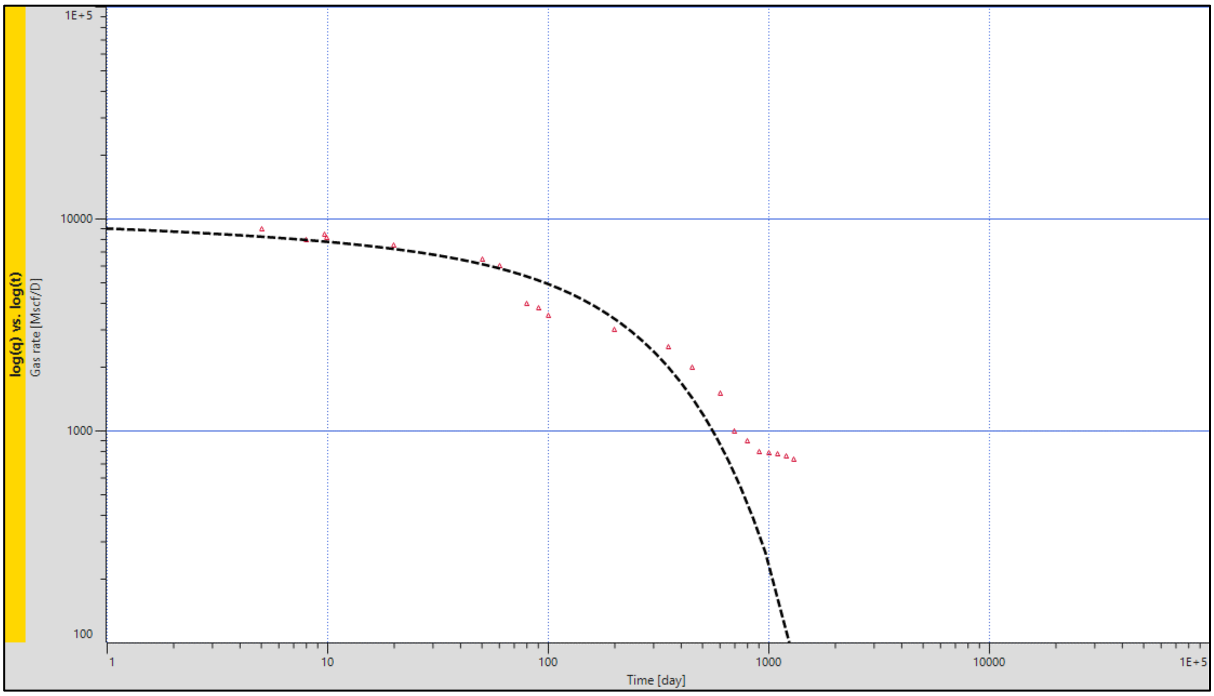


Figure A8: PLE model generated graph from KAPPA-Citrine for the Barnett Shale Well (Tan et al., 2018)

Table A12: PLE model parameters generated by KAPPA-Citrine for the Marcellus Shale Well

Parameters	Power-law exponential [2024/04/12]	
Tmin	0.00000	day
Tmax	12250.0	day
Initial rate	11066.3	Mscf/D
Decline parameter	0.00304357	1/day
Transition time	3134.59	day
n exponent	0.199652	
Abandonment rate	1.91309E-13	Mscf/D
EUR	2.04413	bscf
Relative std. error on rates	112.552	%
Relative std. error on cumulative	5.57909	%
Remaining Recoverable	-0.201714	bscf

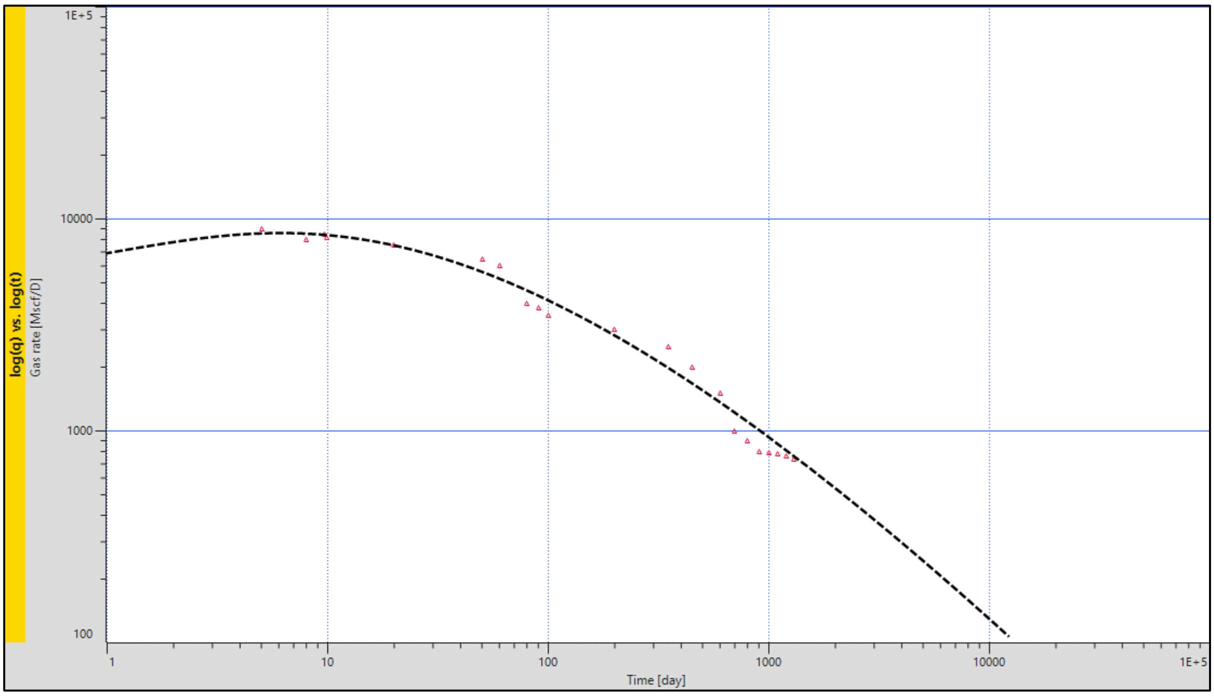


Figure A9: Duong's model generated graph from KAPPA-Citrine for the Barnett Shale Well (Tan et al., 2018)

Table A13: Duong's model parameters generated by KAPPA-Citrine for the Marcellus Shale Well

Parameters	Duong [2024/04/12]	
Tmin	0.00000	day
Tmax	12250.0	day
Initial rate	4968.78	Mscf/D
M exponent	1.20553	
A exponent	1.80467	
Abandonment rate	107.215	Mscf/D
EUR	5.03555	bscf
Relative std. error on rates	11.4295	%
Relative std. error on cumulative	4.68452	%
Remaining Recoverable	2.78971	bscf

A.3. Actual vs. Predicted Data

Table A14: Actual vs. predicted results for the Arps model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
20	260	206
50	240	205
70	200	199
90	160	189
100	140	182
150	110	144
200	108	115
250	105	94
260	101	90
270	100	88
300	90	80
370	80	67
400	70	63
430	60	60
450	55	58
470	53	56
480	52	56
500	51	54
600	50	49
620	49	47
650	45	45
680	47	45
700	42	44
720	40	44
750	36	42

Table A15: Actual vs. predicted results for the Duong's model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
20	260	376
50	240	219
70	200	180
90	160	155
100	140	146
150	110	115
200	108	97
250	105	85
260	101	83
270	100	81
300	90	76
370	80	67
400	70	64
430	60	62
450	55	60
470	53	58
480	52	58
500	51	56
600	50	51
620	49	50
650	45	48
680	47	47
700	42	46
720	40	45
750	36	44

Table A16: Actual vs. predicted results for the PLE model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
20	260	274
50	240	210
70	200	185
90	160	166
100	140	159
150	110	130
200	108	110
250	105	95
260	101	93
270	100	91
300	90	84
370	80	72
400	70	68
430	60	64
450	55	62
470	53	59
480	52	58
500	51	56
600	50	48
620	49	46
650	45	44
680	47	42
700	42	41
720	40	40
750	36	38

Table A17: Actual vs. predicted results for the Arps-Duong-PLE hybrid model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
20	260	254
50	240	217
70	200	196
90	160	178
100	140	169
150	110	135
200	108	110
250	105	92
260	101	89
270	100	86
300	90	78
370	80	66
400	70	62
430	60	59
450	55	57
470	53	56
480	52	55
500	51	54
600	50	49
620	49	49
650	45	48
680	47	47
700	42	47
720	40	47
750	36	46

Table A18: Actual vs. predicted results for the Arps-Duong hybrid model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
20	260	258
50	240	257
70	200	198
90	160	144
100	140	122
150	110	62
200	108	49
250	105	53
260	101	55
270	100	56
300	90	62
370	80	75
400	70	81
430	60	87
450	55	91
470	53	96
480	52	98
500	51	102
600	50	122
620	49	126
650	45	132
680	47	138
700	42	142
720	40	146
750	36	152

Table A19: Actual vs. predicted results for the Arps-PLE hybrid model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
20	260	218
50	240	195
70	200	181
90	160	168
100	140	162
150	110	136
200	108	115
250	105	99
260	101	96
270	100	93
300	90	85
370	80	71
400	70	66
430	60	62
450	55	60
470	53	57
480	52	56
500	51	54
600	50	47
620	49	45
650	45	44
680	47	42
700	42	42
720	40	41
750	36	40

Table A20: Actual vs. predicted results for the Duong-PLE hybrid model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
20	260	207
50	240	199
70	200	193
90	160	188
100	140	179
150	110	171
200	108	166
250	105	138
260	101	132
270	100	130
300	90	124
370	80	119
400	70	110
430	60	102
450	55	83
470	53	75
480	52	72
500	51	69
600	50	55
620	49	41
650	45	28
680	47	25
700	42	19
720	40	14
750	36	6

Table A21: Actual vs. predicted results for the ARIMA model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
20	260	270
50	240	250
70	200	223
90	160	174
100	140	121
150	110	111
200	108	100
250	105	100
260	101	105
270	100	99
300	90	93
370	80	81
400	70	68
430	60	59
450	55	52
470	53	49
480	52	50
500	51	51
600	50	48
620	49	45
650	45	46
680	47	43
700	42	43
720	40	40
750	36	34

Table A22: Actual vs. predicted results for the ANN model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
20	260	280
50	240	214
70	200	181
90	160	155
100	140	145
150	110	113
200	108	104
250	105	102
260	101	101
270	100	100
300	90	96
370	80	78
400	70	69
430	60	61
450	55	56
470	53	53
480	52	51
500	51	49
600	50	48
620	49	48
650	45	48
680	47	46
700	42	43
720	40	39
750	36	30

Table A23: Actual vs. predicted results for the ANN-ARIMA hybrid model – Canon Shale Well (Paryani et al., 2018)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
20	260	286
50	240	221
70	200	188
90	160	162
100	140	151
150	110	120
200	108	111
250	105	108
260	101	108
270	100	107
300	90	102
370	80	84
400	70	75
430	60	67
450	55	63
470	53	59
480	52	58
500	51	56
600	50	55
620	49	55
650	45	54
680	47	52
700	42	49
720	40	45
750	36	37

Table A24: Actual vs. predicted results for the Arps model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
1000	2500	3170
1100	2400	2970
1200	2300	2784
1500	2000	2301
1900	1800	1797
2200	1600	1501
2400	1500	1334
2600	1400	1188
2800	1300	1060
3000	1200	948
3200	1100	849
3400	1000	761
3600	900	684
3800	800	616
4000	700	555
4200	600	502
4400	500	454
4600	400	411
4800	300	373
5000	250	339
6000	150	215
7000	100	142
7500	50	116
10000	40	49
12500	30	24
15000	20	13
17500	10	8

Table A25: Actual vs. predicted results for the Duong's model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
1000	2500	6787
1100	2400	5588
1200	2300	4680
1500	2000	2969
1900	1800	1834
2200	1600	1360
2400	1500	1139
2600	1400	967
2800	1300	832
3000	1200	722
3200	1100	633
3400	1000	560
3600	900	498
3800	800	446
4000	700	402
4200	600	364
4400	500	331
4600	400	302
4800	300	277
5000	250	255
6000	150	176
7000	100	128
7500	50	112
10000	40	62
12500	30	39
15000	20	27
17500	10	20

Table A26: Actual vs. predicted results for the PLE model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
1000	2500	3525
1100	2400	3218
1200	2300	2949
1500	2000	2314
1900	1800	1733
2200	1600	1422
2400	1500	1255
2600	1400	1114
2800	1300	992
3000	1200	888
3200	1100	797
3400	1000	718
3600	900	648
3800	800	587
4000	700	533
4200	600	485
4400	500	442
4600	400	404
4800	300	370
5000	250	339
6000	150	225
7000	100	154
7500	50	129
10000	40	57
12500	30	28
15000	20	14
17500	10	8

Table A27: Actual vs. predicted results for the Arps-Duong-PLE hybrid model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
1000	2500	2558
1100	2400	2453
1200	2300	2352
1500	2000	2073
1900	1800	1750
2200	1600	1541
2400	1500	1414
2600	1400	1298
2800	1300	1191
3000	1200	1092
3200	1100	1000
3400	1000	916
3600	900	839
3800	800	767
4000	700	701
4200	600	640
4400	500	584
4600	400	533
4800	300	485
5000	250	441
6000	150	268
7000	100	152
7500	50	110
10000	40	-10
12500	30	-53
15000	20	-69
17500	10	-75

Table A28: Actual vs. predicted results for the Arps-Duong hybrid model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
1000	2500	2394
1100	2400	2388
1200	2300	2362
1500	2000	2205
1900	1800	1896
2200	1600	1646
2400	1500	1485
2600	1400	1332
2800	1300	1190
3000	1200	1060
3200	1100	943
3400	1000	838
3600	900	745
3800	800	663
4000	700	591
4200	600	529
4400	500	475
4600	400	429
4800	300	390
5000	250	357
6000	150	257
7000	100	226
7500	50	224
10000	40	264
12500	30	326
15000	20	391
17500	10	456

Table A29: Actual vs. predicted results for the Arps-PLE hybrid model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
1000	2500	1940
1100	2400	1892
1200	2300	1845
1500	2000	1711
1900	1800	1548
2200	1600	1436
2400	1500	1366
2600	1400	1299
2800	1300	1235
3000	1200	1175
3200	1100	1118
3400	1000	1063
3600	900	1011
3800	800	962
4000	700	915
4200	600	870
4400	500	828
4600	400	788
4800	300	749
5000	250	713
6000	150	555
7000	100	433
7500	50	383
10000	40	206
12500	30	112
15000	20	62
17500	10	36

Table A30: Actual vs. predicted results for the Duong-PLE hybrid model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
1000	2500	3430
1100	2400	3194
1200	2300	2977
1500	2000	2419
1900	1800	1851
2200	1600	1525
2400	1500	1345
2600	1400	1189
2800	1300	1053
3000	1200	935
3200	1100	832
3400	1000	743
3600	900	664
3800	800	595
4000	700	534
4200	600	481
4400	500	434
4600	400	392
4800	300	355
5000	250	322
6000	150	203
7000	100	134
7500	50	111
10000	40	48
12500	30	25
15000	20	15
17500	10	10

Table A31: Actual vs. predicted results for the ARIMA model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
1000	2500	2509
1100	2400	2414
1200	2300	2305
1500	2000	2204
1900	1800	1787
2200	1600	1565
2400	1500	1521
2600	1400	1399
2800	1300	1297
3000	1200	1214
3200	1100	1103
3400	1000	997
3600	900	911
3800	800	806
4000	700	698
4200	600	608
4400	500	507
4600	400	399
4800	300	306
5000	250	208
6000	150	188
7000	100	81
7500	50	-9
10000	40	35
12500	30	20
15000	20	-27
17500	10	-4

Table A32: Actual vs. predicted results for the ANN model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	**q(t) MSCF/day**	**q(t) MSCF/day**
1000	2500	2478
1100	2400	2387
1200	2300	2299
1500	2000	2055
1900	1800	1774
2200	1600	1595
2400	1500	1488
2600	1400	1388
2800	1300	1293
3000	1200	1201
3200	1100	1108
3400	1000	1011
3600	900	910
3800	800	805
4000	700	697
4200	600	592
4400	500	494
4600	400	406
4800	300	332
5000	250	273
6000	150	138
10000	40	91
12500	30	51
15000	20	5
17500	10	13

Table A33: Actual vs. predicted results for the ANN-ARIMA hybrid model – Marcellus Shale Well (Adekoya, 2009)

	Actual Values	Predicted Values
Time (days)	q(t) MSCF/day	q(t) MSCF/day
1000	2500	2518
1100	2400	2426
1200	2300	2339
1500	2000	2094
1900	1800	1814
2200	1600	1634
2400	1500	1527
2600	1400	1428
2800	1300	1333
3000	1200	1241
3200	1100	1147
3400	1000	1051
3600	900	950
3800	800	844
4000	700	737
4200	600	632
4400	500	533
4600	400	446
4800	300	372
5000	250	313
6000	150	178
7000	100	155
7500	50	150
10000	40	131
12500	30	91
15000	20	45
17500	10	52

A.4. Confidence Intervals

Table A34: Raw data for the 95% confidence interval for the Canon Shale Well

	Actual Values	Predicted Values	Predicted Values		
Time (day)	**Q(t)**	**ANN**	**Arps PLE hybrid decline model**	**LCL**	**UCL**
260	20	29	23	72	119
240	50	42	32	64	128
200	70	52	61	35	156
160	90	78	114	18	210
140	100	110	156	60	252
110	150	200	251	155	347
108	200	208	259	163	355
105	250	221	272	176	368
101	260	239	289	194	385
100	270	243	294	198	390
90	300	295	344	249	440
80	370	356	403	308	499
70	400	426	473	377	568
60	430	506	553	458	649
55	450	550	599	503	695
53	470	568	618	522	714
52	480	577	628	532	724
51	500	587	638	542	734
50	600	596	648	552	744
49	620	606	659	563	755
45	650	626	702	606	798
47	680	646	680	584	776
42	700	677	736	640	832
40	720	698	759	663	855
36	750	741	809	713	905

Table A35: Raw data for the 95% confidence interval for the Marcellus Shale Well

	Actual Values	Predicted Values	Predicted Values		
Time (day)	Q(t)	ANN	Arps PLE hybrid decline model	LCL	UCL
1000	2500	2478	1940	2163	2793
1100	2400	2387	1892	2072	2702
1200	2300	2299	1845	1984	2614
1500	2000	2055	1711	1740	2370
1900	1800	1774	1548	1459	2089
2200	1600	1595	1436	1280	1910
2400	1500	1488	1366	1172	1803
2600	1400	1388	1299	1073	1703
2800	1300	1293	1235	978	1609
3000	1200	1201	1175	886	1516
3200	1100	1108	1118	793	1423
3400	1000	1011	1063	696	1326
3600	900	910	1011	595	1225
3800	800	805	962	489	1120
4000	700	697	915	382	1012
4200	600	592	870	277	907
4400	500	494	828	178	809
4600	400	406	788	91	721
4800	300	332	749	17	647
5000	250	273	713	42	588
6000	150	138	555	177	454
10000	40	91	206	224	407
12500	30	51	112	264	367
15000	20	5	62	310	320
17500	10	13	36	302	328

The Evaluation and Sensitivity of Decline Curve Modelling

Prinisha Manda * and Diakanua Bavon Nkazi

Oil and Gas Production and Processing Research Unit, School of Chemical and Metallurgical Engineering, University of the Witwatersrand, Johannesburg 2000; South Africa; Diakanua.Nkazi@wits.ac.za

* Correspondence: 690765@students.wits.ac.za; Tel.: +27-834117636

Received: 28 March 2020; Accepted: 26 April 2020; Published: date

Abstract: The development of prediction tools for production performance and the lifespan of shale gas reservoirs has been a focus for petroleum engineers. Several decline curve models have been developed and compared with data from shale gas production. To accurately forecast the estimated ultimate recovery for shale gas reservoirs, consistent and accurate decline curve modelling is required. In this paper, the current decline curve models are evaluated using the goodness of fit as a measure of accuracy with field data. The evaluation found that there are advantages to using the current DCA models. However, they also have limitations associated with them that have to be addressed. Based on the accuracy assessment conducted on the different models, it appears that the Stretched Exponential Decline Model (SEDM) and the Logistic Growth Model (LGM), followed by the Extended Exponential Decline Model (EEDM), the Power Law Exponential Model (PLE), the Duong's Model, and lastly, the Arps Hyperbolic Decline Model, provide the best fit with production data.

Keywords: valuation; shale gas reservoirs (SGR); decline curve models; decline curve analysis (DCA); estimated ultimate recovery (EUR)

---

1. Introduction

In recent years, shale gas reservoirs (SGR) or unconventional reservoirs have steadily become the main bases of natural gas production around the world [1]. Wang [2] notes that shales and sediments are the richest sedimentary rocks in the Earth's crust and, according to recent activities, shale gas will constitute the largest component in gas production globally, as conventional reservoirs continue to decrease. It is further mentioned by Wang [2] that SGR, unlike conventional reservoirs, tend to be more costly to develop and require special tools to enable the gas to be produced at a cost-effective rate, due to their extremely low matrix permeability and porosity [3]. Accordingly, the modelling of shale gas production and its decline is essential to predict how fast the gas can be produced and turned into revenue from each well, as well as modelling of the feasibility of producing natural gas from operated shale plays from a cost perspective [2].

Currently, the oldest and most commonly used tool for the modelling of shale gas production is the rate versus time decline curve estimation, due to its ease. Current efforts in decline curve analysis (DCA) have been concentrating on a computer statistical approach, the basic objective being to arrive at a distinctive "unbiased" interpretation [2]. In recent years, several DCA models have been suggested and compared with previous shale gas production figures, prior to being used on more reservoirs [4]. This paper focuses on the evaluation and sensitivity of the current DCA models and proposes a new hybrid model to be investigated in SGR decline analysis. The main ideas are (a) to characterise and evaluate the current decline curve models used to explain shale gas reservoir forecasting and (b) to use the goodness-of-fit regression test to assess the sensitivity of the decline curve models in (a).

2. Overview of Shale Gas Production

Yuan et al. [5] specified that due to the rise in energy demand and the decrease in conventional oil and gas production, shale gas has received increasing attention worldwide. Shale gas presently makes up more than 20% of the drilled gas production in the United States (US) [6]. The manipulation of shale gas is land-based and generally requires a sizeable quantity of wells to achieve beneficial recovery rates [6]. Nwaobi and Anandarajah [7] explained that shale gas reservoir production and viability have been investigated globally, but that progress has been slow due to a number of concerns, one of which is a precise production forecast. Nwaobi and Anandarajah [7] went on to define that the quantity of shale gas reserve that can be recovered is the estimated ultimate recovery (EUR) for the petroleum industry. The EUR is a key factor for stakeholders and policymakers in evaluating petroleum resources [7].

Shale gas production has been vital in providing the US, which was formerly a natural gas importer, with the ability to export natural gas [5]. This development aided the US to effectively ensure its energy security and decrease its carbon emissions considerably. Canada has become the second nation to attain viable exploitation of its shale gas reserves. The positive development of shale gas in Canada has resuscitated the nation's natural gas production, which had formerly experienced a rapid decrease [5].

3. Characteristics and Production Behavior of Shale Gas

Shale gas reservoirs possess the characteristics such as ultra-low permeability, no trap mechanism, and the gas is tightly absorbed to the rock particle, which is the opposite of a conventional reservoir [8]. Hydraulic fracturing is often used in reservoirs with low permeability that are not able to reach economic production rates [8]. This is very different in character to the naturally fractured reservoirs that are classified as having dual porosity [8]. There are four different flow regimes that can occur in a hydraulically fractured reservoir and several flow periods can exist during the life cycle of a shale gas well [8,9]. These consist of fracture linear flow, fracture boundary flow, matrix linear flow, and lastly matrix boundary flow [10]. Joshi [10] explained the different flow regimes for shale gas reservoirs as follows:

- **Fracture/Early Linear Flow**: A transient flow regime that occurs when the production flow is linear to the single fractures. This flow regime governs the known life of most shale wells. A negative half slope on a log–log plot of rate versus time can be used to differentiate this linear flow.
- **Fracture Boundary Flow**: Follows after a certain period of production when an interference occurs i.e., from linear to simulated reservoir volume (SRV). Many of the existing horizontal shale wells have not experienced this regime, but some of the newer wells with huge fracture treatments have been observing this regime early. This can be observed on a log–log plot by deviation from a −1/2 slope line on a log–log plot of rate versus time.
- **Matrix Linear Flow**: When production from the matrix, beyond the SRV starts to govern the production, a linear type of flow will be seen. This regime will probably not be observed in the economic life of the well. Comparable to fracture linear flow, this regime can be observed using a negative half slope line on a log–log plot of rate versus time.
- **Matrix Boundary Flow**: After the outer matrix transient has reached the drainage boundaries of the well, a deviation from the negative half slope, corresponding to matrix linear flow, will be observed. This deviation is equivalent to matrix boundary flow. Similarly to the matrix, linear flow will probably not be observed.

4. Overview of Decline Curve Models

Consistently forecasting the long-term production performance of shale (unconventional) reservoirs has been a challenge [11]. The petroleum industry requires simple, useful, and speedy means of predicting production and assessing reserves. Hence, DCA has been an attractive alternative in contrast to other methods [11]. Due to the relative ease of DCA, it is considered the most used method in the industry [11]. The current DCA models will be evaluated based on their characteristics, strengths, weaknesses, and sensitivity to production data.

4.1. Arps Decline Curve and the Modified Hyperbolic Decline Model (MHD)

Arps decline curve analysis is the most commonly used method of estimating ultimate recoverable reserves and future performance [12]. Paryani et al. [13] recommended this to be a reliable history match (even with b > 1); also due to its simplicity. The model process is based on vital assumptions: that past operating conditions will remain unaffected, a well is produced at or near capacity, and that the well's drainage remains constant and is produced at a constant bottom-hole pressure [14]. Notably, the Arps model is only applicable in pseudo-steady flows when the flow regime transfers from linear flows to boundary-dominated flows (BDF) [15]. This indicates that the Arps equations are not applicable to the production forecasting of the entire decline process of horizontal wells in low-permeability reservoirs [16]. The Arps decline curve analysis can be summarised into three types: exponential Equation (1), hyperbolic Equation (2), and harmonic Equation (3) [17,18].

$$q = qie^{-Dt} \tag{1}$$

$$q = \frac{q_i}{(1 + bD_it)^{\frac{1}{b}}} \tag{2}$$

$$q = \frac{qi}{1 + D_i} \tag{3}$$

where q is the flow rate in STB/day or Mscf/day, q_i is the initial flow rate in STB/day or Mscf/day, D is the decline constant while D_i is the initial decline constant, which are both measured in days-1, and b is the decline exponent.

The most commonly employed hyperbolic form of Arps decline Equation (2) is used for shale reservoirs. The hyperbolic decline equation is suitable to use due to the "best fit" it provides for the long transient linear-flow regime observed in shale gas wells with b values greater than unity [18]. The model results in post-production overestimation, due to the decrease in the decline rate with production time. Due to the overestimation, Robertson et al. [19] suggested a revised version of the hyperbolic decline model for shale gas production decline. The equation is given as:

$$q = \frac{q_i}{(1 + nD_it)^{1/n}} \;\; (D > D_{lim}) \tag{4}$$

$$q = q_i \; exp(-D_{lim}t) \;\; (D \leq D_{lim}) \tag{5}$$

where q is the production rate in m^3/d or STB/day, D_{lim} is the decline rate in d^{-1}, and n is the time exponent. They suggested that the hyperbolic decline model sometimes yields unrealistically high reserve estimates. They made an assumption that the rate of decline starts at 30% of flow and usually declines in a hyperbolic way [19]. This modified model considers when the hyperbolic decline in the early life of a well transfers to exponential decline in the late life [19]. The switching process can be determined by applying computer programs. The switching point is when the decline rate is smaller than a certain limit (usually 5%) [19]. The MHD model addresses the overestimation limitation of EUR; however, it is still unable to determine D_{lim} for production data [15].

To test the behavior of the Arps hyperbolic model and the modified version shown in Figure 1, a semi log plot (log q versus t) illustrates the sensitivity of the models to various estimated field data. The R^2 values denote the goodness of fit or the degree of linear correlation, which is a measure of the level of association of a group of actual observations to the model's forecasts [20]. As observed from the regression lines for the various data, the resulting fit appears to capture the trend in the data well. Arps fits Data 1 and 2 fairly, similarly for the MHD. However, the method matches the other cases poorly because it cannot model multiple flow regimes. In the case of the MHD model, there is a shift in the curves downward, which results in a change in the R^2 value. Upon closer inspection of the EUR values for both models, which are shown in Table 1, it is evident that the MHD model corrects for the overestimation of the Arps model.

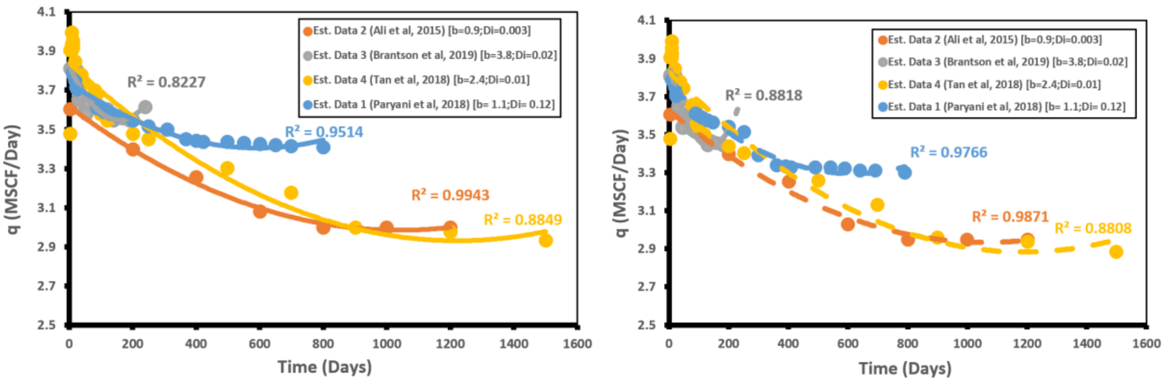


Figure 1. Data sensitivity using the Arps Hyperbolic Decline Model [4,13,14,21].

Table 1. Summary of estimated ultimate recovery (EUR). MHD: Modified Hyperbolic Decline Model.

EUR (bscf)	Data 1	Data 2	Data 3	Data 4
Arps Model	0.31	20.52	18.13	5.21
MHD	0.18	4.13	13.18	4.18

4.2. Power Law Exponential Model (PLE)

Ilk et al. [22] presented the PLE, which is an extension of the exponential Arps formula for the decline degree in shale reservoirs. This model was developed precisely for SGR and approximates the rate of decline with a power law decline. The PLE model matches production data in both the transient and boundary-dominated regions, without being hypersensitive to remaining reserve estimates [23]. Seshadri and Mattar [24] presented that the PLE model can model transient radial and linear flows, while Kanfar and Wattenbarger [25] proved that the model is reliable for linear flow, bilinear flow followed by linear flow, and linear flow followed by BDF, or bilinear flow followed by linear flow and finished with BDF flow. Vanorsdale [26] deduced that when the flow regime changes throughout the initial 10 years of the well, the PLE model will yield a very optimistic recovery. The model characterizes the decline rate by infinite time, D_∞ which is defined as a "loss ratio" (which is assumed to be constant from Arp) [16]. The production rate is derived as follows:

$$\frac{q}{dq/dt} = -b \tag{6}$$

$$b = D_{\infty} + D_i t^{-(1-\hat{n})} \tag{7}$$

where dq/dt is the slope, D_∞ is the decline rate over a long-term period, and $\hat{n}$ is the time exponent. By substituting the above equations, the production rate is obtained:

$$q(t) = \hat{q}_i e^{[-D_\infty t - \hat{D}_i t^{\hat{n}}]}. \tag{8}$$

In this model, there are four unknown variables: $\hat{q}_i, \widehat{D}_{i,}\ D_{\infty}$ *and* $\hat{n}$, which result in several degrees of freedom and may be clumsy to use or solve [27]. According to Johnson et al. [28], the $D\infty$ parameter is difficult to determine. However, there are advantages to this model in that the extra variables permit for both transient and boundary flow, and the equation for production rate seems comparable to the Arps exponential equation [13]. With the PLE model (Figure 2), which uses a log–log plot (log q versus log t) to test the sensitivity of the data, the resulting fit appears to capture the trend in the data better, compared to the Arps Hyperbolic Model. This model fits Data 1, 2, 3, and 4 fairly accurately. This can be attributed to the PLE model, matching production data in both transient and boundary-dominated regions.

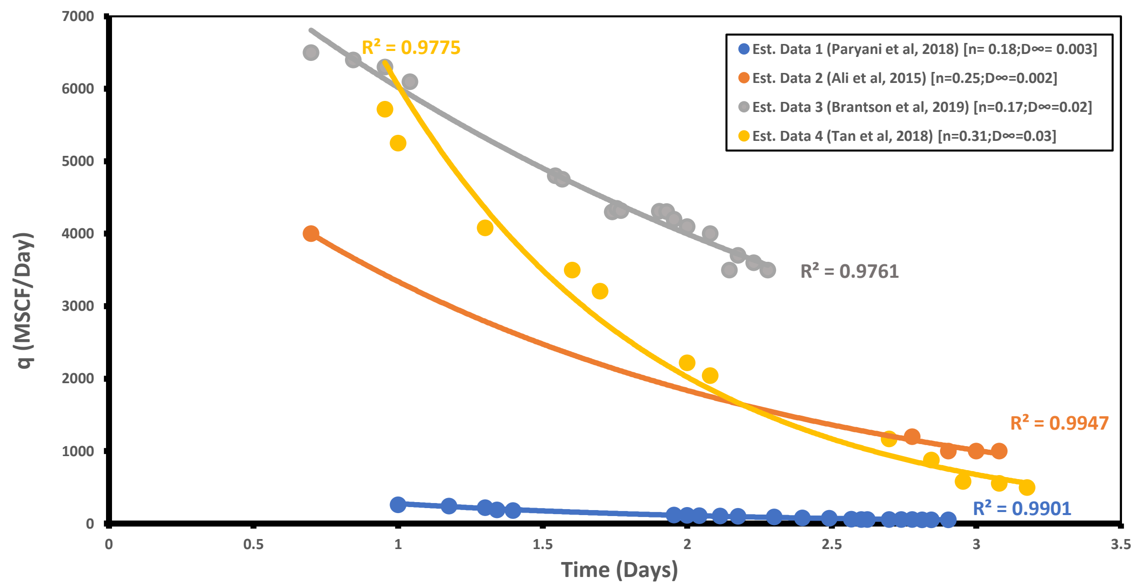


Figure 2. Data sensitivity using the Power Law Exponential Model [4,13,14,21].

4.3. Stretched Exponential Decline Model (SEDM)

Valkó, Valkó, and Lee [29,30] applied the SEDM in shale wells, which is an empirical method different from Arps equations, as it describes the decline trend of production data obtained from unconventional reservoirs. It was developed to fit transient flow regimes [10,25]. The significant advantages of the model are the bounded nature of estimated ultimate recovery (EUR) without limits on time or rate, and the straight-line behavior of a recovery potential expression [30]. The model differs from other models since it does have a basis in physics and is directed by a major differential equation [14]. It is used to model aftershock decay rates [31]. The production rate declines with time, according to the following equations:

$$\frac{dq}{dt} = -n\left(\frac{t^n}{\tau}\right)\frac{q}{t} \tag{9}$$

$$q = q_i \exp\left[-\left(\frac{t}{\tau}\right)^n\right] \tag{10}$$

$$Q = \frac{qi^n}{n}\left\{\Gamma\left[\frac{1}{n}\right] - \Gamma\left[\frac{1}{n}.\left(\frac{t}{\tau}\right)^n\right]\right\} \tag{11}$$

$$\text{EUR} = \frac{qi^{\tau}}{n}\,\tau\left[\frac{1}{n}\right]. \tag{12}$$

This method defines a characteristic number of periods, τ, and a dimensionless exponent, n, of the ratio of time, t. It also uses observed cumulative production along with theoretical cumulative production, derived from the integral of the rate-time equation to estimate remaining technically recoverable volumes. Equation (10) appears similar to the PLE model; however, it differs, as it does not rely on a single interpretation of parameters. Instead, it uses two-parameter gamma functions [29]. In addition, there is no single τ and n parameters, but instead, a sum of multiple exponential declines, which follows the fat tail distribution [30]. Stretched Exponential Decline Model (SEDM) requires an iterative process to determine the value of the parameter, n. The model can only estimate the recoverable volumes with an abandonment rate of zero, as opposed to commercial volumes with economic cut-off rates and has not been widely used [32]. However, Can et al. [32] showed that in tight formations where transient flow period is extremely long, SEDM has been successful in modeling rate-time behavior and provides more realistic reserve estimates compared to Arps decline relations.

Testing the behavior of the SEDM, Figure 3, which is a plot of production rate versus the cumulative production (*q* versus *Q*) to test the sensitivity of the data, the resulting fit appears to capture the trend in the data poorly. The SEDM method fits all cases inaccurately (lower R^2 values). This is due to the SEDM model's transient flow rather than boundary-dominated flow and requirement for a sufficiently long production time (usually >36 months) to accurately estimate the parameters τ and *n* [33].

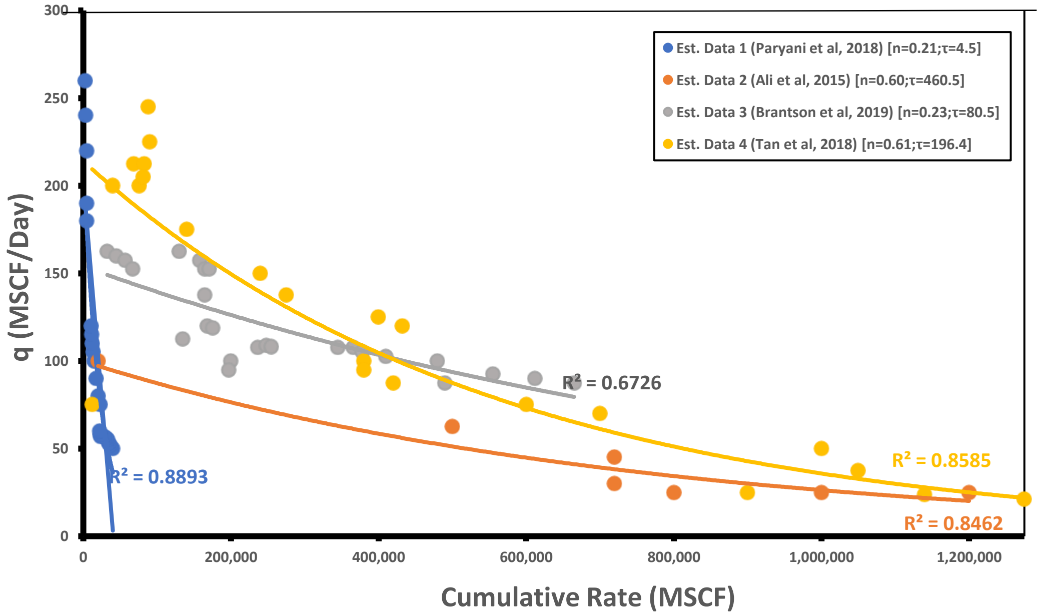


Figure 3. Data sensitivity using the Stretched Exponential Decline Model [4,13,14,21].

4.4. The Extended Exponential Model (EEDM)

Zang et al. [11] presented a renewed experimental method, the EEDM, as a simple formula to forecast shale oil and gas well performance. They proposed a mechanism of "growing drainage volume" to conceptualize and model the performance of shale wells. This model combines the exponential decline equation proposed by Fetkovich et al. [34] Equation (13) with the derived empirical Equation (14). The EEDM includes both transient and BDF flow in a single equation, and it can match the historical data with a smooth curve throughout the transition period from transient to BDF flow regimes. Furthermore, the model is simple and can easily be applied [11]. It is also able to project future production by fitting all of the historical production data from the beginning of the production decline.

Paryani et al. [13] stated that the model contains two decline constants and a decline exponent. particularly noteworthy, is that the production data fits using a smooth curve through the whole flow systems [16]. The advantage of the model is that both early and late production profiles can be captured once β_e and β_l have been calibrated using the production data [11]. However, as parameter β_l has an incomplete influence on the curve fitting, it is therefore fixed.

$$q = q_i e^{-at} \tag{13}$$

$$a = \beta_l + \beta_e \tag{14}$$

where a is the nominal decline rate, β_l is the late-life period constant, and β_e is the early period constant. Combining Equations (13) and (14) and taking the logarithm of each side, the equation below (the exponential decline equation) is obtained.

$$\frac{In\frac{q}{q_o}}{t} = \beta_l + \beta_e e^{-t^n} \tag{15}$$

where q_o is the initial production rate in m^3/s. Using the EEDM (Figure 4), which is a plot of $-ln\frac{q}{qo}/t$ versus t to test the sensitivity of the data, the resulting fit appears to also capture the trend in the data poorly. The method fits all cases inaccurately (lower R^2 values). This type of method is best for forecasting short-term trends in the absence of recurring variations. Hence, the EEDM would only be accurate when a realistic amount of stability between the past and future is assumed.

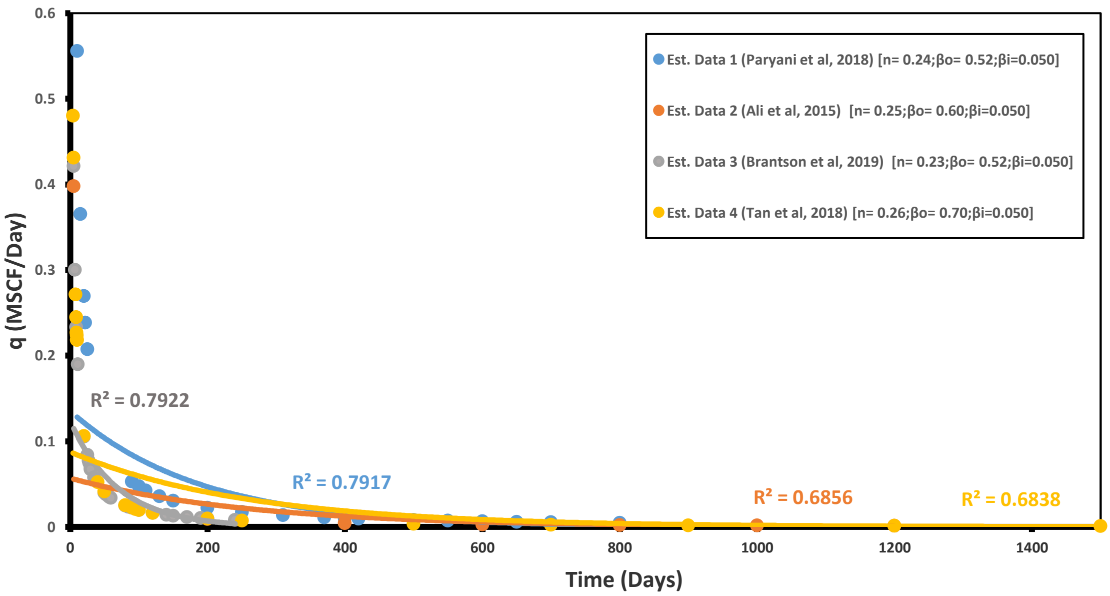


Figure 4. Data sensitivity using the Extended Exponential Model [4,13,14,21].

4.5. Doung's Decline Model

Duong [35] presented an unconventional rate decline method to evaluate the performance of shale gas wells that does not depend on fracture types. The model assumes linear or near-linear flow, as indicated by a log–log plot of rate over cumulative production versus time, which yielded a straight-line tendency [36]. The rate is calculated in the model using the following equation [27]:

$$q(t) = q_i t(a,m) + q_\infty \tag{16}$$

where t (a,m) is the time constant in 1/s, and q_∞ is the production rate at infinite time in m^3/s. The cumulative production and time constant is calculated as:

$$G_p = \frac{qt(a,m)}{at^{-m}} \tag{17}$$

$$t(a,m) = t^{-m}\exp(\frac{a}{1-m}(t^{1-m} - 1)) \tag{18}$$

where G_p is the cumulative gas production in Bcf and m is the slope.

Paryani et al. [13] indicated the key restrictions of the model: if the well is closed for extended periods, a proper rate initialization against pressure is required to obtain precise values of parameters a and m and, secondly; that in the event of water breakthrough, there is a sudden decrease in the decline rate, this causes an increase in the values of the a and m parameters. Vanorsdale [26], similar to the case of the PLE model, also indicated that the Duong model will yield a very optimistic recovery when the flow regime changes throughout the initial 10 years. He went on to indicate that the model may provide a conservative recovery estimate in vertical, non-hydraulic fractured classical shale wells [26]. However, Lee et al. [36] indicated that the Duong's model appears to fit field data from various shale plays quite well and provides an effective alternative to the Arps hyperbolic model. With the Duong's model (Figure 5), which uses a log–log linear plot (log q versus log t) to test the sensitivity of the data, the resulting fit appears to capture the trend in the data well. The method fits Data 1, 2, and 4 fairly accurately. For Case 3, the method fits the data poorly with a lower R^2 value of 0.8371. The model probably provides a good fit because it was specifically developed for unconventional reservoirs with very low permeability.

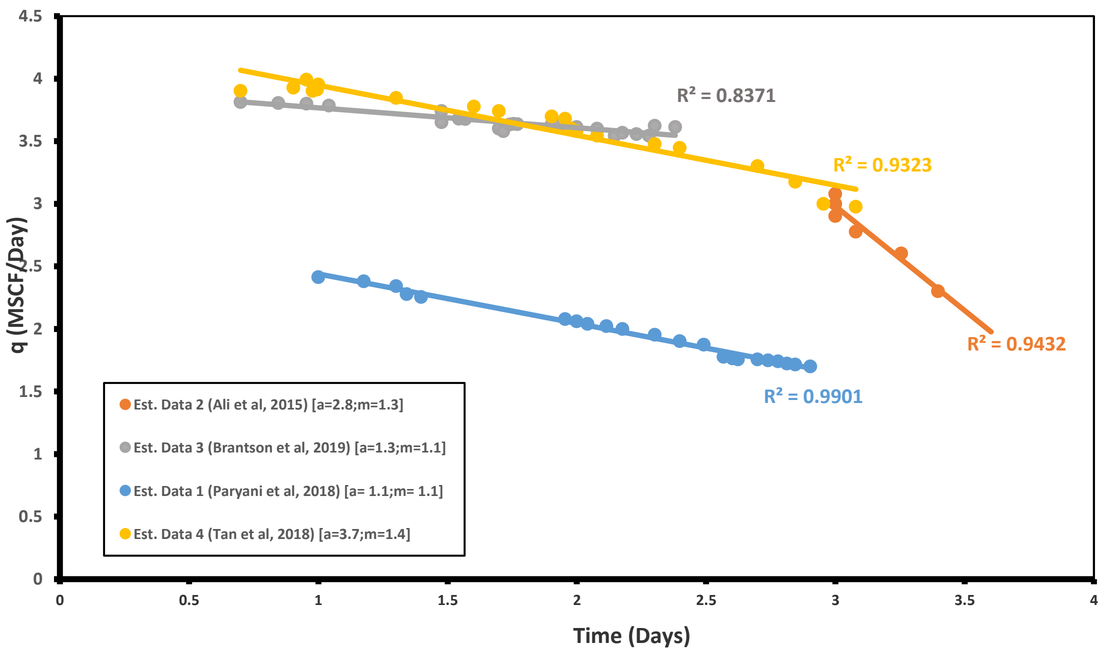


Figure 5. Data sensitivity using Duong's Decline Model [4,13,14,21].

4.6. Logistic Growth Model (LGM)

Logistic Growth Models developed belong to a group of mathematical models used to forecast growth in numerous applications [36] and were previously used to model population growth [37,38]. It was developed to forecast reservoirs with extremely low permeability [27]. LGM is very flexible and confident in modelling long transient boundary-dominated performances of unconventional reservoirs [16]. The model incorporates known physical volumetric quantities of oil and gas into the forecast, to constrain the reserve estimate to a reasonable quantity. LGM is capable of trending existing production data and providing reasonable forecasts of future production. The logistic growth model does not extrapolate to non-physical values [38]. Tsoularis and Wallace [39] discussed a development in this regard by Verhulst [40], who considered that for the population model, a steady population would consequently possess a saturation level characteristic, typically called the carrying capacity, K, which forms a numerical upper bound on the growth size. In order to include this limiting characteristic, they introduced the logistic growth equation as an extension to the exponential model [39]. Zhang et al. [1] adopted this model for SGR with very low permeability and developed the LGM as an empirical method to forecast gas production. The LGM can be represented as follows:

$$q(t) = \frac{dQ}{dt} = \frac{Knbt^{n-1}}{(a + t^{n})^2} \tag{19}$$

where K is the carrying capacity.

The main benefit of the LGM is that the reserve estimate is inhibited by the parameter K as well as the production rate, which terminates at infinite time [1]. The main assumption in this model is that the whole reservoir can be drained by a single well over a suitably long period and requires the approximation of at least two parameters, or parameters as per the available well information [4,7]. Figure 6, a plot of production rate versus time (q versus t), illustrates the sensitivity of the model to various estimated field data. As observed from the regression lines for the various data, the resulting fit appears to capture the trend in the data well. The LGM fits Data 1 and 2 fairly. However, the method matches the other cases poorly, as indicated by the lower R^2 values. This could be attributed to the data size, which is too small to yield an accurate fit, since the underlying principle of this model is population growth, which stipulates that growth is only possible up to a certain size.

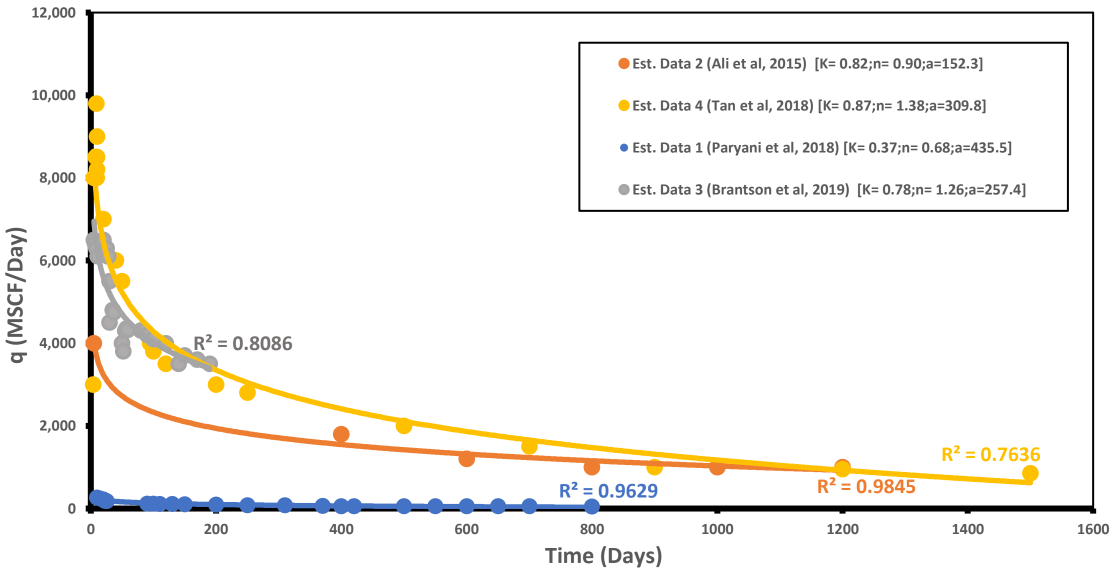


Figure 6. Data sensitivity using the Logistic Growth Model [4,13,14,21].

4.7. Autoregressive Intergrated Moving Average (ARIMA) and Neutral Network Models (NNM) (Hybrid Model)

The accuracy of time series forecasting is challenging for scientists [41]. Time series data often comprises linear as well as non-linear components [42]. In some cases, linear-based approaches might be more suitable than non-linear ones due to the data characteristics. The hybrid method is a combination of ARIMA and the neural network method. According to Faruk [42], hybrid methods have a higher degree of accuracy than neural networks. ARIMA can recognize time-series patterns well but not non-linear data patterns. On the other hand, neural networks only handle non-linear data. Therefore, hybrid models combine the advantages of ARIMA with respect to linear modelling and neural networks in terms of non-linear edge modelling [43]. Notwithstanding, in some circumstances, the single model approach can outperform hybrid models [41]

Mathematically, time-series data can be expressed as a combination of linear and non-linear components [44]:

$$Y_t = L_t + N_t \tag{20}$$

where Y_t shows the time-series data, L_t indicates the linear components, and the non-linear components are represented by N_t.

Mathematically, the neural network model for residual of n input nodes can be expressed as the following:

$$e_t = f(e_{t-1} + e_{t-2}, \dots, e_{t-n}) \tag{21}$$

where f is a non-linear function that is specified by the neural network. With regard to the results of the prediction error of N_t, the combination forecast using the hybrid method can be expressed as:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t. \tag{22}$$

There has been limited work conducted using this model for shale gas reservoirs. Hence, the next step would be to investigate this model for shale gas reservoirs.

To summarize all eight DCA models for an easy reference of readers, Table 2 lists the name of each model, its DCA equation, the characteristic, strength, weakness, and lastly the related references.

Table 2. Summary of decline curve analysis (DCA) models. BDF: boundary-dominated flows, SGR: shale gas reservoirs.

No	Model	Equation	Production Behaviour	Strength	Weakness	Reference
1	Arps Hyperbolic Decline	$q = \frac{q_i}{(1 + bD_i t)^{\frac{1}{b}}}$	linear to BDF flow	reliable and simple to use	post-production overestimation	[12–18]
2	Modified Hyperbolic Curve	$q = \frac{q_i}{(1 + nD_i t)^{1/n}} \quad (D > D_{lim})$ $q = q_i\, exp(-D_{lim} t) \quad (D \leq D_{lim})$	transient and BDF flow	addresses the overestimation limitation of EUR	still unable to determine D_{lim} for production data	[15,19]
3	Power Law Exponential Decline	$q(t) = \hat{q}_i e^{[-D_\infty t - \hat{D}_i t^{\hat{n}}]}$	transient and BDF flow	developed precisely for SGR	four unknown variables to solve	[13,16,20,23–27]
4	Stretched Exponential Decline	$q = q_i\, ex\, p\left[-\left(\frac{t}{\tau}\right)^n\right]$	transient flow	bounded nature of EUR and straight-line behaviour of recovery potential expression	requires sufficiently long production times	[10,14,26,28–32]
5	The Extended Exponential Model	$\frac{\mathrm{In}\frac{q}{q_o}}{t} = \beta_l + \beta_e e^{-t^n}$	transient and BDF flow	both early and late production profiles can be captured	parameter β_l has an incomplete influence on the curve fitting and is therefore fixed	[11,13,16,33]
6	Duong's Decline	$t(a,m) = t^{-m} exp(\frac{a}{1-m}(t^{1-m} - 1))$	linear or near-linear flow	appears to fit field data from various shale plays	extended periods, a proper rate initialization against pressure is required, and in the event of water breakthrough, a and m increases	[13,20,27,34,35]
7	Logistic Growth	$q(t) = \frac{dQ}{dt} = \frac{Knbt^{n-1}}{(a + t^n)^2}$	long transient boundary-dominated	reserve estimate is inhibited by K as well as the production rate, which terminates at infinite time	growth is only possible up to a certain size	[1,16,20,35–39]
8	Hybrid Model	$\emptyset\,(B)(1-B)^d Y_t = \theta(B)\varepsilon_t$	linear and non-linear	high degree of accuracy	approach can be found to not be fit all types of data	[40–43]

5. Accuracy of Current Decline Curve Models with Field Data

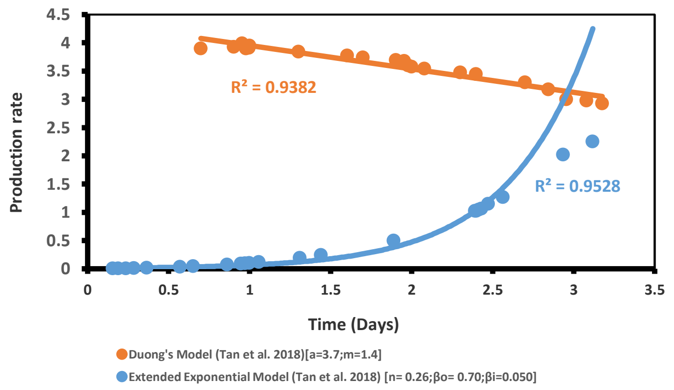
Yuhu et al. [15] discussed comparisons of EURs with five types of decline models from single-well production data. They explained that according to the prediction results, the highest predicted EUR was gained by the hyperbolic decline model, followed by the Modified Hyperbolic Model (MHD), Duong's Model, PLE and, lastly, the EDM. Hu et al. [27] conferred production data for wells with a production time greater than 10 years. Therefore, the PLE decline model was recommended for multiple flows. It was also pointed out that the hyperbolic decline model predicted higher estimates of reserves than the PLE decline model. Another study that they reviewed recommended the MHD rather than the PLE decline model, which in their view was complicated.

It is noted that the differences in EURs with different decline models decrease with an increase in production time [45]. On the other hand, prediction consistency increases with an increase in production time. Based on this distinctive production data, the order of predicted EURs from high to low was through the hyperbolic decline model, the MHD, the PLE decline model, and the EDM respectively [45]. The predicted EURs decreased with an increase of production time for the hyperbolic decline and the modified hyperbolic decline model. The predicted EURs increase with an increase of production time for the PLE decline and the EDM model [45]. Currently, the applicability of these different decline models is uncertain. The general trend found in Yuhu et al's paper was that the hyperbolic decline model overestimates the production and that the other decline models will still have to be investigated for reliability and accuracy [45].

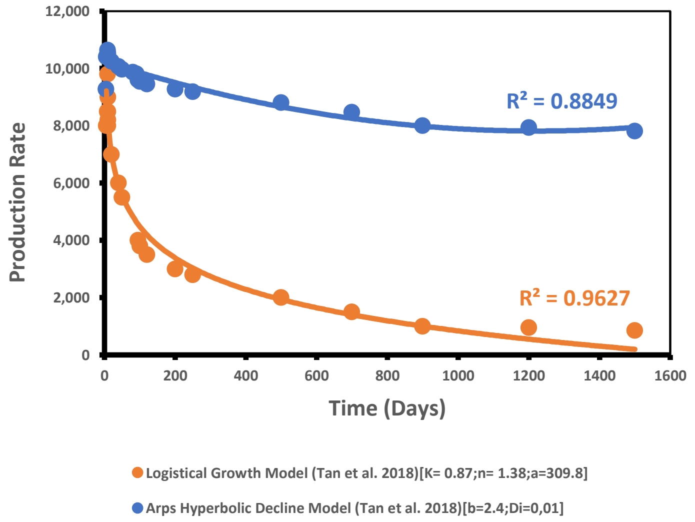
Guo et al. [46] investigated shale gas wells in the Barnett shale play, where they found that from the results of goodness of fit, the hyperbolic curve fits well for both the aggregate and individual shale gas wells. On the other hand, Kenomore et al. [47] in their production decline study of the Barnett shale found that either the Arps hyperbolic or Duong's model can be used only if the historical data exceeds 10 months. They used root mean square error (RMSE) analysis and the results indicated that the Arps hyperbolic model showed better forecasting compared to the Duong's model for the top three longest production histories. Zhang et al. [1] concurred with the findings of the Duong's model, noting that it is more accurate for linear flow and bilinear flow; however, if the production history is shorter than 18 months, this model provides unreliable results for EUR. In most circumstances, the Duong's model overestimates the total EUR. Harris [48], in his research study of the Elm Coulee field production data, found that the Duong method produces the most optimistic forecasts followed by the Arps model with 5% minimum decline, and then the SEPD model. Shah [49] in his research developed new methods of combining the SEPD and Arps hyperbolic equation, the Duong's with the Arps hyperbolic equation, and the Arps super hyperbolic combined with the Arps hyperbolic decline equation. He found that the SEPD and Arps hyperbolic equation gave the most conservative results of all the methods in the study, even if there was insufficient data available. This equation can also work without enough boundary-dominated flow (BDF) data being available.

Hu et al. [27] studied DCA techniques for the Eagle Ford and Austin Chalk reservoirs. They found that in the case of the Eagle Ford reservoir, the MHD and the Duong's model provided the highest EUR estimations and the two lowest matching errors, while the PLE decline model with $D_\infty \neq 0$ produced the lowest EUR estimates with the highest matching errors in all cases. In another study, according to the results of goodness of fit (R^2 and N-RMSE), the hyperbolic model fits well with aggregated well data and with individual wells [1]. Furthermore, Hu et al. [27] explained that the LGM and PLE model with $D_\infty = 0$ gave production projections that were neither too positive nor too traditional with modest matching errors. Therefore, they recommend both the MHD and Duong's model for this reservoir. However, Zhang et al. [11] developed the EEDM and verified their model using field data from Eagle Ford. They found this model to be more rigorous in that it included the effects of interference among adjacent fractures, variable permeability, and discontinuous pressure distribution, which are difficult to capture and model with other DCA methods [11]. In the case of the Austin Chalk reservoirs, all DCA methods resulted in fairly similar EUR forecasts and matching errors; hence, any method can be used [27]. Figure 7, which uses estimated production data versus time values, indicates that using the R^2

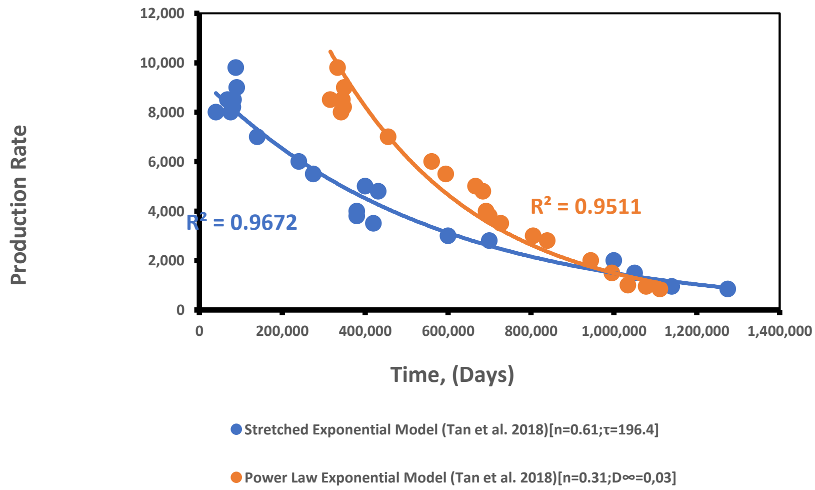
values as a goodness of fit to determine the accuracy of the different decline models, the SEDM, followed by the LGM, EEDM, PLE, Duong's decline model and, lastly, the hyperbolic decline model would predict the EUR accurately.



(**a**)



(**b**)



(**c**)

Figure 7. Estimated production data to determine goodness of fit for accuracy of the different decline models (a) Duong's Model vs EEDM; (b) LGM vs Arps Hyperbolic Model and (c) SEDM vs PLE [4].

During their case study analysis, Paryani et al. [13] found that the LGM, PLE, and Duong's models overcame Arps limitations to a certain degree. The PLE model always predicted the lowest forecasts of all the models, with the most conservative production forecasting and reserve estimation. Duong's model performed the best for longer when less noisy production data was available; however, erratic EUR was observed, which indicates that this model requires further improvements [13]. The LGM gave reasonable EUR estimates when compared to the Arps model. There was an 81% fit of the wells' past production rate and cumulative production. The LGM also appears most effective at historically matching past production and predicting finite reasonable EUR. However, Tan et al. [4] found that due to the constraints of *K* and the vanishing production rate at infinity time, the LGM provides a finite estimate of EUR. They also determined by using normalized and logarithmic rate-time residuals that the limitations of the Arps model are overcome and accuracy improves in cases of unconventional reservoirs.

6. Conclusions

Shale gas reservoirs have become an essential source for providing natural gas globally and the process of hydraulic fracking has been used in the extraction of shale gas. During the fracking process, there are different flow regimes, which occur during the life cycle of SGRs; these being fracture linear flow, fracture boundary flow, matrix linear flow, and matrix boundary flow. They are significant because they impact both the production and decline behavior of SGRs.

Based on previous studies, it was found that the Arps hyperbolic decline, the MHD and Duong's models provided the best fit with production data. However, contrary to the reviewed studies when estimated production data was used in the evaluation process for the basis of this paper, using the goodness-of-fit technique, the PLE and Duong's decline models aligned best with the production data, compared to the other models.

It is evident from the accuracy assessment that decline curve modelling impacts the EUR of SGRs, and it was observed that all decline models yield a different EUR result, which is either over or under-estimated. Studies have revealed that the production time significantly impacts the EUR, depending on which decline model is being used. When each model was assessed for accuracy, once again using the goodness-of-fit technique, the results indicated that the SEDM, followed by the LGM, EEDM, PLE, Duong's decline model and, lastly, the hyperbolic decline model aligns with the production data.

It is evident from the decline curve evaluation that there are advantages in using the current DCA models; however, they also have limitations associated with them, which have to be addressed. Therefore, the next step will be to evaluate the use of the hybrid model in evaluating the decline of SGR.

Author Contributions: Conceptualization, P.M. and D.N.; methodology, P.M. and D.N.; validation, P.M. and D.N.; formal analysis, P.M.; investigation, P.M.; data curation, P.M.; writing—original draft preparation, P.M.; writing—review and editing, P.M.; visualization, P.M.; supervision, D.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, X.; Wang, X.; Hou, X.; Xu, W. Rate decline analysis of vertically fractured wells in shale gas reservoirs. *Energies* **2017**, *10*, 1602.
2. Wang, H. What factors control shale gas production and production decline trend in fractured systems: A comprehensive analysis and investigation, *SPE J.* **2017**, *22*, 562–581.
3. Xu, B.; Haghighi, M.; Li, X.; Cooke, D. Development of new type curves for production analysis in naturally fractured shale gas/tight gas reservoirs. *J. Pet. Sci. Eng.* **2013**, *105*, 107–115.

4. Tan, L.; Zuo, L.; Wang, B. Methods of decline curve analysis for shale gas reservoirs. *Energies* **2018**, *11*, 552.
5. Yuan, J.; Luo, D.; Feng, L. A review of the technical and economic evaluation techniques for shale gas development. *Appl. Energy* **2015**, *148*, 49–65.
6. Knudsen, B.R.; Foss, B.; Whitson, C.H.; Conn, A.R. Target-rate tracking for shale-gas multi-well pads by scheduled shut-ins. *IFAC Proc. Vol.* **2012**, *45*, 107–113.
7. Nwaobi, U.; Anandarajah, G.A. Critical Review of Shale Gas Production Analysis and Forecast Methods. *Saudi J. Eng. Technol. (SJEAT)* **2018**, *3*, 276–285.
8. Adekoya, F. Production Decline Analysis of Horizontal Well in Gas Shale Reservoirs. Master's Thesis, West Virginia University, Morgantown, WV, USA, 2009.
9. Nelson, P.H. Pore-throat sizes in sandstones, tight sandstones, and shales. *AAPG Bull.* **2009**, *93*, 329–340.
10. Joshi, K.J. Comparison of Various Deterministic Forecasting Techniques in Shale Gas Reservoirs with Emphasis on the Duong Method. Ph.D. Thesis, Texas A&M University, TX, USA, 2012.
11. Zhang, H.; Rietz, D.; Cagle, A.; Cocco, M.; Lee, J. Extended exponential decline curve analysis. *J. Nat. Gas Sci. Eng.* **2016**, *36*, 402–413.
12. Boah, E.A.; Borsah, A.A.; Brantson, E.T. Decline Curve Analysis and Production Forecast Studies for Oil Well Performance Prediction: A Case Study of Reservoir X. *Int. J. Eng. Sci. (IJES)* **2018**, *7*, 56–67.
13. Paryani M.; Ahmadi M.; Awoleke O.; Hanks C. Decline Curve Analysis: A Comparative Study of Proposed Models Using Improved Residual Functions. *J. Pet. Environ. Biotechnol.* **2018**, *9*, 362.
14. Ali, T.A.; Sheng, J.J. Production decline models: A comparison study. In Proceedings of the SPE Eastern Regional Meeting, Morgantown, WV, USA, 13–15 October 2015.
15. Yuhu, B.; Guihua, C.; Bingxiang, X.; Ruyong, F.; Ling, C. Comparison of typical curve models for shale gas production decline prediction. *China Pet. Explor.* **2016**, *21*, 96–102.
16. Li, P.; Hao, M.; Hu, J.; Ru, Z.; Li, Z. A new production decline model for horizontal wells in low-permeability reservoirs. *J. Pet. Sci. Eng.* **2018**, *171*, 340–352.
17. Arps, J.J. Analysis of decline curves. *Trans. AIME* **1945**, *160*, 228–247.
18. Qu, Z.; Lin, J.E. (Eds.) *Proceedings of the International Field Exploration and Development Conference 2017*; Springer: Singapore, 2018.
19. Robertson, S. *Generalised Hyperbolic Equation*; Society of Petroleum Engineers: Richardson, TX, USA, 1988.
20. Bagozzi, R.P.; Yi, Y. On the evaluation of structural equation models. *J. Acad. Mark. Sci.* **1988**, *16*, 74–94.
21. Brantson, E.T.; Ju, B.; Ziggah, Y.Y.; Akwensi, P.H.; Sun, Y.; Wu, D.; Addo, B.J. Forecasting of Horizontal Gas Well Production Decline in Unconventional Reservoirs using Productivity, Soft Computing and Swarm Intelligence Models. *Nat. Resour. Res.* **2019**, *28*, 717–756.
22. Ilk, D.; Rushing, J.A.; Perego, A.D.; Blasingame, T.A. Exponential vs hyperbolic decline in tight gas sands: Understanding the origin and implications for reserve estimates using Arps decline curves. In Proceedings of the SPE Annual Technical Conference and Exhibition, Denver, CO, USA, 21–24 September 2008.
23. McNeil, R.; Jeje, O.; Renaud, A. Application of the power law loss-ratio method of decline analysis. In Proceedings of the Canadian International Petroleum Conference, Calgary, AB, Canada, 16–18 June 2009.
24. Seshadri, J.N.; Mattar, L. Comparison of power law and modified hyperbolic decline methods. In Proceedings of the Canadian Unconventional Resources and International Petroleum Conference, Calgary, AB, Canada, 19–21 October 2010.
25. Kanfar, M.S.; Wattenbarger, R.A. Comparison of Empirical Decline Curve Methods for Shale Wells. In Proceedings of the SPE Canadian Unconventional Resources Conferences, Calgary, AB, Canada, 30 October–1 November 2012.
26. Vanorsdale, C.R. Production decline analysis lessons from classic shale gas wells. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 30 September–2 October 2013.
27. Hu, Y.; Weijermars, R.; Zuo, L., Yu, W. Benchmarking EUR estimates for hydraulically fractured wells with and without fracture hits using various DCA methods. *J. Pet. Sci. Eng.* **2018**, *162*, 617–632.
28. Johnson, N.L.; Currie, S.M.; Ilk, D.; Blasingame, T.A. A Simple methodology for direct estimation of gas-in-place and reserves using rate-time data. In Proceedings of the SPE Rocky Mountain Petroleum Technology Conference, Denver, CO, USA, 14–16 April 2009.
29. Valko, P.P. Assigning value to stimulation in the Barnett Shale: A simultaneous analysis of 7000 plus production hystories and well completion records. In Proceedings of the SPE Hydraulic Fracturing Technology Conference, The Woodlands, TX, USA, 19–21 January 2009.

30. Valkó, P.P.; Lee, W.J. A better way to forecast production from unconventional gas wells. In Proceedings of the SPE Annual Technical Conference and Exhibition, Florence, Italy, 19–22 September 2010.
31. Kisslinger, C. The stretched exponential function as an alternative model for aftershock decay rate. *J. Geophys. Res.: Solid Earth* **1993**, *98*, 1913–1921.
32. Can, B.; Kabir, C.S. Probabilistic performance forecasting for unconventional reservoirs with stretched-exponential model. In Proceedings of the North American Unconventional Gas Conference and Exhibition, The Woodlands, TX, USA, 14–16 June 2011.
33. Zhou, L.Z.; Selim, H.M. Application of the fractional advection-dispersion equation in porous media. *Soil Sci. Soc. Am. J.* **2003**, *67*, 1079–1084.
34. Fetkovich, M.J. Decline curve analysis using type curves. *J. Pet. Technol.* **1980**, *32*, 1065–1077.
35. Duong, A.N. Rate-decline analysis for fracture-dominated shale reservoirs. *SPE Reserv. Eval. Eng.* **2011**, *14*, 377–387.
36. Lee, K.S.; Kim, T.H. *Integrative Understanding of Shale Gas Reservoirs*; Springer: Heidelberg, Germany, 2016.
37. Clark, A.J. Decline Curve Analysis in Unconventional Resource Plays Using Logistic Growth Models. Ph.D. Thesis, The University of Texas at Austin, Austin, TX, USA, 2011.
38. Clark, A.J.; Lake, L.W.; Patzek, T.W. Production forecasting with logistic growth models. In Proceedings of the SPE Annual Technical Conference and Exhibition, Denver, CO, USA, 30 October–2 November 2011.
39. Tsoularis, A.; Wallace, J. Analysis of logistic growth models. *Math. Biosci.* **2002**, *179*, 21–55.
40. Bacaër, N. Verhulst and the logistic equation (1838). In *A Short History of Mathematical Population Dynamics*; Springer: London, UK, 2011; pp. 35–39
41. Taskaya-Temizel, T.; Ahmad, K. Are ARIMA neural network hybrids better than single models? In Proceedings of the International Joint Conference on Neural Networks (IJCNN 2005), Montreal, QC, Canada, 31 July–4 August 2005.
42. Faruk, D.Ö. A Hybrid Neural Network and ARIMA Model for Water Quality Time Series Prediction. *Eng. Appl. Artif. Intell.* **2010**, *23*, 586–594.
43. Cybenko, G. Approximation by Superpositions of a Sigmoidal Function. *Math. Control Signals Syst.* **1989**, *2*, 303–314.
44. Dhini, A.; Riefqi, M.; Puspasari, M.A. Forecasting analysis of consumer goods demand using neural networks and ARIMA. *Int. J. Technol.* **2015**, *6*, 872–880.
45. Wachtmeister, H.; Lund, L.; Aleklett, K.; Höök, M. Production decline curves of tight oil wells in eagle ford shale. *Nat. Resour. Res.* **2017**, *26*, 365–377.
46. Guo, K.; Zhang, B.; Wachtmeister, H.; Aleklett, K.; Höök, M. Characteristic Production Decline Patterns for Shale Gas Wells in Barnett. *Int. J. Sustain. Future Hum. Secur.* **2017**, *5*, 11–20.
47. Kenomore, M.; Hassan, M.; Malakooti, R.; Dhakal, H.; Shah, A. Shale gas production decline trend over time in the Barnett Shale. *J. Pet. Sci. Eng.* **2018**, *165*, 691–710.
48. Harris, S.C. A Study of Decline Curve Analysis in the Elm Coulee Field. Ph.D. Thesis, Texas A&M University, Texas, USA, 2013.
49. Shah, S. Development of New Decline Model for Shale Oil Reserves. Ph.D. Thesis, University of Houston, Houston, TX, USA, 2013.
50. Makinde, I.; Lee, W.J. Forecasting production of liquid rich shale (LRS) reservoirs using simple models. *J. Pet. Sci. Eng.* **2017**, *157*, 461–481.
51. Paryani, M.; Awoleke, O.O.; Ahmadi, M.; Hanks, C.; Barry, R. Approximate Bayesian Computation for Probabilistic Decline-Curve Analysis in Unconventional Reservoirs. *SPE Reserv. Eval. Eng.* **2017**, *20*, 478–485.
52. Mishra, S. A new approach to reserves estimation in shale gas reservoirs using multiple decline curve analysis models. In Proceedings of the SPE Eastern Regional Meeting, Lexington, KY, USA, 3–5 October 2012.
53. Duong, A.N. An unconventional rate decline approach for tight and fracture-dominated gas wells. In Proceedings of the Canadian Unconventional Resources and International Petroleum Conference, Calgary, AB, Canada, 19–21 October 2001.
54. Ilk, D.; Rushing, J.A.; Blasingame, T.A. Integration of production analysis and rate-time analysis via parametric correlations--theoretical considerations and practical applications. In Proceedings of the SPE Hydraulic Fracturing Technology Conference, The Woodlands, TX, USA, 24–26 January 2011.
55. Lee, W.J.; Sidle, R. Gas-reserves estimation in resource plays. *SPE Econ. Manag.* **2010**, *2*, 86–91.

56. Zhang, H.; Cocco, M.; Rietz, D.; Cagle, A.; Lee, J. An empirical extended exponential decline curve for shale reservoirs. In Proceedings of the SPE Annual Technical Conference and Exhibition, Houston, TX, USA, 28–30 September 2015.
57. Zhang, H.E.; Nelson, E.; Olds, D.; Rietz, D.; Lee, W.J. Effective Applications of Extended Exponential Decline Curve Analysis to both Conventional and Unconventional Reservoirs. In Proceedings of the SPE Annual Technical Conference and Exhibition, Dubai, UAE, 26–28 September 2016.

Production Decline Prediction of Shale Gas using Hybrid Models

P. Manda and D.B. Nkazi

Abstract

Hybrid models have frequently been used for shale gas production decline prediction by manipulating the unique strength of each of the known decline models. The use of a combination of models provides a more precise predicting model for forecasting time series data as compared to an individual model. In this study, the forecasting performance of decline curve hybrid models and ANN-ARIMA hybrid models are evaluated and compared with Arps', Duong's, the Power Law Exponential Decline, Autoregressive Integrated Moving Average (ARIMA) and Artificial Neutral Network (ANN) models, respectively. The variable used to assess the models was the respective flow rate, q(t) monitored over a period of time (T). The results have shown that the single model approach can outperform hybrid models. The average deviation of the two best models indicates a central tendency of the production data around the mean. Subsequently, the spread in the data between the actual and predicted values is found to be less. It can thus be concluded that the ARIMA and ANN models have the best forecasting accuracy for production decline in shale gas compared to the other models.

Keywords

Shale gas decline forecasting, Arps' decline model, Duong's decline model, PLE decline model; ARIMA, ANN and hybrid models

1. Introduction

Rate-time decline curve extrapolation is one of the oldest and most commonly used tools by a petroleum engineer. Results obtained for a well are subject to a wide range of alternate interpretations, mostly as a function of the experience and objectives of the evaluator. Recent efforts in the area of decline curve analysis (DCA) have been directed towards a purely computerised statistical approach, its basic objective being to arrive at a unique "unbiased" interpretation [1]. In the past few decades, several DCA models have been proposed and benchmarked with commercial reservoir simulators or shale gas production data before being applied to more shale gas reservoirs (SGRs) [2].

Numerous studies have highlighted the importance of DCA models, however, there are limitations with these models. Analysis conducted using these techniques for the prediction and estimation of reservoirs in shale well production have highlighted shortcomings in the models [3]. These shortcomings include underestimation, finite and overestimation of the estimated ultimate recovery (EUR) of reserves. Taking these facts into consideration, the scope exists for developing improved models which address these shortcomings.

1.1 Production Decline Models

The Arps decline model is inaccurate within the transient flow regime (TFR) and the Duong model is inaccurate within the boundary dominated flow (BDF). Although the Power Law Exponential (PLE) model incorporates both these flow regimes and was specifically developed for SGRs, the model has its own shortcomings. Hence, the scope to develop a new decline model or a new method to predict more accurately the recovery of SGRs. Accordingly, the approach would be to combine the above-mentioned methods i.e. to evaluate the hybrid decline curve models. As the PLE and Duong's models model the transient flow well and because the Arps model is widely used for BDF, the new approach combines the methods to achieve the objectives and eliminate the shortcomings of the stand-alone models. In this paper, the combination of different models, or hybrid models as they are commonly known, will be investigated.

Hybrid models have frequently been used for prediction by manipulating the unique strength of each of the models [4]. The use of a combination of models provides a more precise predicting model for forecasting time series data as compared to an individual model [5]. The results from studies have indicated that hybrid models have higher prediction accuracy for one-step and multi-step forward forecasts and various hybrid models have been used for obtaining accurate prediction [5; 6].

The evaluation of the forecasting performance of decline curve hybrid models and ARIMA-ANN hybrid models is essential, and these models should be compared with Arps', Duong's, the Power Law Exponential Decline, Autoregressive Integrated Moving Average (ARIMA) and Artificial Neutral Network (ANN) models for accurate prediction of production decline in shale gas.

1.2 Hybrid Models and ANN-ARIMA Hybrid Models

In the literature, hybrid methods are considered to yield better results [7]. The accuracy of time series forecasting is challenging for scientists [7]. Time series data often comprise linear as well as non-linear components [8]. In some cases, linear-based approaches might be more suitable than non-linear approaches due to the data characteristics. The use of hybrid models, which combine DCA models, is a new approach and there is minimal literature covering this aspect. However, as mentioned, the known approach to the hybrid method is a combination of the ARIMA and ANN method.

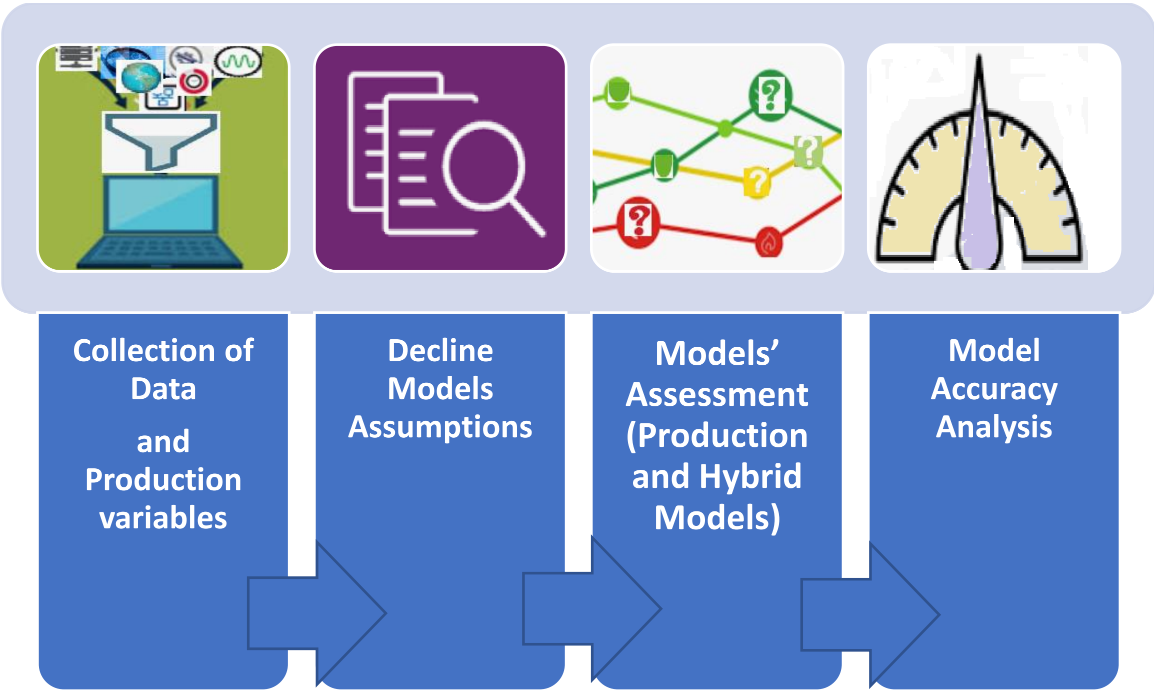
According to Faruk [8], hybrid methods have a higher degree of accuracy than neural networks. ARIMA is able to recognise time-series patterns well except non-linear data patterns. On the other hand, neural networks only handle non-linear data. Therefore, hybrid models combine the advantages of ARIMA with respect to linear modelling and neural networks in terms of non-linear edge modelling [9]. Ayub and Jafri (2020) [10] in their paper highlighted that the combined model has improved forecasting accuracy as compared to when the models are used individually. Notwithstanding this, in some circumstances the single model approach can outperform hybrid models [8]. Babu et al. (2014) [5] explored ARIMA and ANN as a new hybrid model for better prediction of time series. Their results preferred the use of the hybrid model compared to the individual ARIMA and ANN models.

The ARIMA processes follow a stochastic behaviour used to analyse time series [11] and is mostly used to predict demand. The application of the ARIMA methodology for the study of time series analysis was developed by Box and Jenkins [11]. The Box–Jenkins methodology includes three iterative steps of model identification, parameter estimation and diagnostic checking [12]. This three-step model building process is typically repeated several times until a satisfactory model is finally selected and can then be used for prediction purposes [12]. In an ARIMA model, the future value of a variable is assumed to be a linear function of several past observations and random errors [11]. During the past decades, researchers have been focusing more on linear models due to their simplicity in comprehension and application (Fattah et al. 2018). A disadvantage of the classical ARIMA methodology is that it requires a large number of observations to determine the best fit model for a data series [13].

The ANN model, on the other hand, has found increasing consideration in forecasting theory, leading to successful applications in various forecasting domains including economics, business, financial and many more [14]. One advantage of neural networks compared to other non-linear models is their universal model, which is capable of predicting fairly extensive functions with a high degree of accuracy. No assumptions are required for neural networks, thus neural networks conform to the characteristics of the data [15]. However, there are disadvantages associated with this model such as constructing the forecasting model, the selection of the network architecture and the data pre-processing techniques which apply to the time series data [16;17].

This investigation uses different hybrid models in forecasting production decline and evaluating the hybrid models for improved forecasting accuracy of time series by using the unique strengths of the models. The experimental results used are based on the study of shale gas production data obtained from a previous study done by Paryani et al. [3].

2. Methodology



2.1 Collection of Data

The variable used in this investigation is flow rate, q(t) in STB/day, monitored over a period of time (T) in days. The estimated data was extracted from the research conducted by Paryani et al. (2018), who obtained the data from the Cannon Well located in Karnes County evaluated over a two-year period. Kappa Citrine and JMP software are used for simulation of the DCA, hybrid DCA, ARIMA, ANN and ANN-ARIMA hybrid models respectively.

2.2 Production Behaviour

2.2.1. Arps' Decline Curve Model

Arps' decline curve analysis is the most commonly used method of estimating ultimate recoverable reserves and future performance [18]. Paryani et al. [3] attribute this to reliable history match (even with b > 1) and its simplicity. The model process is based on the following vital assumptions: that past operating conditions will remain unaffected; that a well is produced at or near capacity; and that the well's drainage remains constant and is produced at a constant bottom-hole pressure [19]. Notably, the Arps model is only applicable in pseudo-steady flows when the flow regime transfers from linear flows to boundary-dominated flows (BDF) [20]. This indicates that the Arps equations are not applicable to the production forecasting of the entire decline process of horizontal wells in low-permeability reservoirs [21]. The most commonly employed hyperbolic form of Arps' decline equation [1] is used for shale reservoirs. The hyperbolic decline equation is suitable to use due to the "best fit" that it provides for the long transient linear-flow regime observed in shale gas wells with b values greater than unity [22].

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

where q is the flow rate in STB/day or Mscf/day, q_i is the initial flow rate in STB/day or Mscf/day, D_i is the initial decline constant, which is measured in days^{-1}, and b is the decline exponent.

Table 1: Summary of the Arps model behaviour, assumptions, condition and parameters

Production Behaviour	Assumptions	Condition	Parameters
Boundary Dominated Flow (BDF)	Decline parameter, b, defines the decline behaviour	$0 < b < 1$	$b = 1.10$ $D_i = 0.12$

2.2.2. Duong's Decline Curve Model

Duong [23] presented an unconventional rate decline method to evaluate the performance of shale gas wells that does not depend on the fracture types. The model assumes linear or near-linear flow, as indicated by a log–log plot of rate over cumulative production versus time, which yielded a straight-line tendency [24]. The rate is calculated in the model using the following equation [2]:

$$q(t) = q_i t(a, m) + q_\infty \tag{2}$$

where $t\,(a,m)$ is the time constant in 1/s, and q_∞ is the production rate at infinite time in m^3/s.

Table 2: Summary of the Duong model's behaviour, assumptions, condition and parameters

Production Behaviour	Assumptions	Condition	Parameters
Transient Flow Regime (TFR)	Very low permeability and long periods of transient flow	b > 1	q_i = 361.24 a = 1.07 m = 1.10

2.2.3. Power Law Exponential Decline Model (PLE)

Ilk et al. [25] presented the PLE, which is an extension of the exponential Arps formula for the decline degree in shale reservoirs. This model was developed precisely for SGR and approximates the rate of decline with a power law decline. The PLE model matches production data in both the transient and boundary-dominated regions without being hypersensitive to remaining reserve estimates [26]. Seshadri and Mattar [27] presented that the PLE model can model transient radial and linear flows, while Kanfar and Wattenbarger [28] proved that the model is reliable for linear flow, bilinear flow followed by linear flow, and linear flow followed by BDF, or bilinear flow followed by linear flow and finished with BDF flow. Vanorsdale [29] deduced that when the flow regime changes throughout the initial 10 years of the well, the PLE model will yield a very optimistic recovery. The model characterises the decline rate by infinite time, D_∞ which is defined as a "loss ratio" (which is assumed to be constant from Arp) [30]. The production rate is derived as follows:

$$\frac{q}{dq/dt} = -b \tag{3}$$

$$b = D_{\infty} + D_i t^{-(1-\hat{n})} \tag{4}$$

where dq/dt is the slope, D_∞ is the decline rate over a long-term period, and $\hat{n}$ is the time exponent. By substituting the above equations, the production rate is obtained:

$$q(t) = \hat{q}_i e^{[-D_\infty t - \hat{D}_i t^{\hat{n}}]}. \tag{5}$$

Table 3: Summary of the PLE model behaviour, assumptions, condition and parameters.

Production Behaviour	Assumptions	Condition	Parameters
BDF and TFR	Approximates the rate of decline with a power law decline	b changes with time	n = 0.182 D_i = 0.268

2.2.4. The Arps'-Duong's-Power Law Models Hybrid Model

The first proposed method incorporates the three DCA models, namely Arps', Duong's and PLE models. The Arps model only considers BDF while Duong's and PLE models consider TFR. The PLE model also considers BDF and has been specifically developed for SGRs. Hence, by combining the three models the limitations from each is presumed to be minimised or eliminated. The equation is given as:

$$\frac{qt}{qi} = t\,(-D_{\infty} - D_i\hat{n}) - In\frac{b+1}{b} \tag{6}$$

where q_t is the flow rate in STB/day or Mscf/day, q_i is the initial flow rate in STB/day or Mscf/day, t is the time in days, $D\infty$ the decline rate over a long-term period, while D_i is the initial decline constant, which are both measured in days $^{-1}$, $\hat{n}$ is the time exponent and *b* is the decline exponent.

Table 4: Summary of the Arps-Duong-Power Law hybrid model behaviour, assumptions, condition and parameters.

Production Behaviour	Assumptions	Condition	Parameters
BDF and TFR	Approximates the rate of decline with an exponential decline	0 > b >1	n = 0.182 D_i = 0.194 b = 1.10

2.2.5. The Arps-Duong Hybrid Model

The second proposed model incorporates the two developed DCA models. Arps' model only considers BDF while Duong's considers TFR, hence both these flow regimes will be taken into account when combining these two models. The equation is given as:

$$qt = [\frac{qt}{t}][1 + bD_i]^{-\frac{1}{b}} \tag{7}$$

where qt is the flow rate in STB/day or Mscf/day, t is the time in days, D_i is the initial decline constant, which is measured in days $^{-1}$ and b is the decline exponent.

Table 5: Summary of the Arps-Duong hybrid model behaviour, assumptions, condition and parameters.

Production Behaviour	Assumptions	Condition	Parameters
BDF and TFR	Approximates the rate of decline with a mechanistic growth decline	0 > b >1	D_i = 0.194 b = 1.10

2.2.6. The Arps-Power Law Exponential Hybrid Model

The third proposed model incorporates the Arps and PLE models. These models consider BDF and TFR flows. Since the PLE model was developed specifically for SGRs, it would be advantageous to evaluate these two models combined due to both being simple equations to use. The equation is given as:

$$t[-D_\infty - D_i\hat{n}] = \frac{-\frac{1}{b}}{qt} In(1 + bD_i) \tag{8}$$

where qt is the flow rate in STB/day or Mscf/day, t is the time in days, $D\infty$ the decline rate over a long-term period and D_i the initial decline constant, which are both measured in days^{-1}, $\hat{n}$ is the time exponent and b is the decline exponent.

Table 6: Summary of the Arps-Power Law Exponential hybrid model behaviour, assumptions, condition and parameters.

Production Behaviour	Assumptions	Condition	Parameters
BDF and TFR	Approximates the rate of decline with a logistic decline	0 > b >1	n = 0.182 D_i = 0.194 b = 1.10

2.2.7. The Duong-Power Law Exponential Hybrid Models

The fourth proposed model incorporates the Duong and PLE models. These models both consider TFR. The equation is given as:

$$\frac{In\ qt}{qm} = t\ [-D_\infty - D_i \hat{n}] \tag{9}$$

where qt is the flow rate in STB/day or Mscf/day, t is the time in days, $D\infty$ the decline rate over a long-term period and D_i the initial decline constant, which are both measured in $days^{-1}$ and $\hat{n}$ is the time exponent. q_m is the flow rate at slope m in m^3/s.

Table 7: Summary of the Duong-Power Law Exponential hybrid model behaviour, assumptions, condition and parameters.

Production Behaviour	Assumptions	Condition	Parameters
BDF and TFR	Approximates the rate of decline with a mechanistic growth decline	0 > b >1	n = 0.182 D_i = 0.194 q_m = 7.12

2.2.8. Autoregressive integrated Moving Average (ARIMA) Model

As mentioned earlier in the paper, the ARIMA processes follow a stochastic behaviour used to analyse time series [11] and are mostly used to predict production demand. The model is labelled as an ARIMA model (p, d, q), where: -

4. p is the number of autoregressive terms;
5. d is the number of differences; and
6. q is the number of moving averages.

According to Ayub and Jafri (2020) [10], the best ARIMA model is determined according to criteria as follows:

- Relatively small BIC
- Maximum adjusted R^2

2.2.8.1. The Autoregressive Process

This process assumes that Y_t is a linear function of the preceding values and is given by equation (5).

$$Y_t = \propto_1 Y_{t-1} + \varepsilon_t \tag{10}$$

Generally, each observation consists of a random component i.e. a random shock, ε and a linear combination of the previous observations. $\propto_1$ in the equation is the self-regression coefficient.

2.2.8.2. The Integrated Process

The integrated process is the archetype of non-stationary series. A differentiation of order 1 assumes that the difference between two successive values of Y is constant. An integrated process is defined by equation (6).

$$Y_t = Y_{t-1} + \varepsilon_t \tag{11}$$

where the random perturbation ε_t is a white noise.

2.2.8.3. The Moving Average Process

The moving average process is a linear combination of the current disturbance with one or more previous perturbations. The moving average order indicates the number of previous periods embedded in the current value. Thus, a moving average is defined by equation (7).

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{12}$$

In order to evaluate the best fit for the ARIMA model, a number of scenarios were evaluated and the ARIMA scenario (2,1,2) was selected to give the best forecast values, due to having the lowest MSE of 4.82, a low BIC of 168,23 and highest adjusted R^2 of 0,979. **Table 2** indicates the best results for the ARIMA model, which are highlighted in bold.

Table 2: Statistical results for the different p,d,q for the ARIMA model

ARIMA	BIC	MSE	Adjusted R^2
(0,0,0)	8,63	46.91	0.000
(1,1,1)	6,19	5.86	0.974
(1,2,1)	9.42	5.84	0.958
(1,3,1)	6,69	6.35	0.899
(2,1,1)	8,25	5.08	0.974

(2,1,2)	**8,23**	**4.82**	**0.979**

2.9. Artificial Neutral Network (ANN) Model

The model consists of three interconnected layers: the input layer, the hidden layer, and the output layer. The basic unit of any ANN is the neuron or node (processor). Each node is able to sum many inputs x1, x2,..., x3 whether these inputs are from a database or from other nodes, with each input modified by an adjustable connection weight [14]. The relationship that occurs in the output and input layers follows equation (8).

$$Y_t = \propto_0 + \sum_{j-1}^{q} \propto_j g\left(\beta_0 j + \sum_{l=1}^{p} \beta_i j Y_t - i\right) + \varepsilon_t \quad (13)$$

where $\propto_j$ (j = 1,2,3, ..., q) and $\beta_i j$ (i = 1,2,3, ..., p; j = 1,2,3, ..., q) are the parameters of the model (often called the weights), p is the number of input points (input nodes), and q is the number of hidden nodes. The activation function used in the hidden layer is the logistic sigmoid function and the linear function is the output layer.

To choose the best algorithm for the model, the number of hidden nodes and layers are changed. The accuracy can also be increased by increasing the number of nodes and layers [31]. In the case of this study, a univariate input layer and four nodes as shown in **Figure 1** gave the best model.

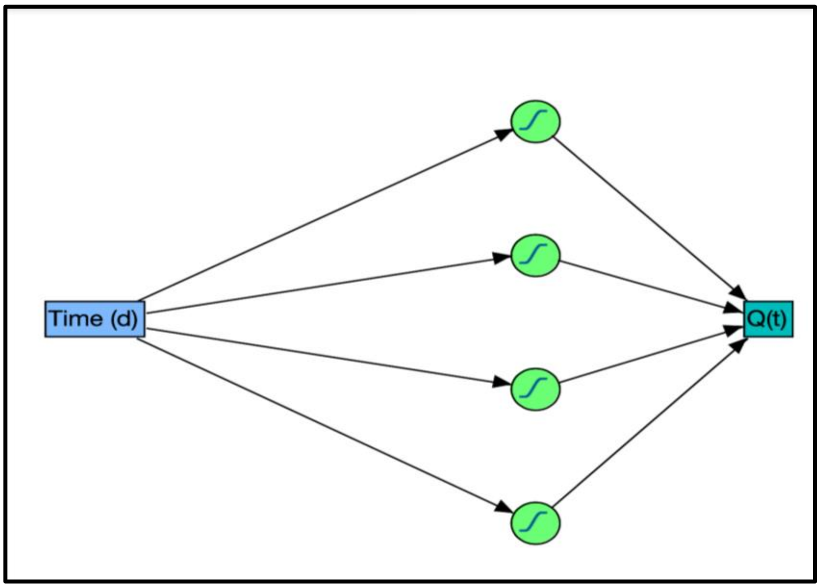


Figure 1: Univariate Artificial Neutral Network obtained from JMP

2.10. ANN-ARIMA Hybrid Model

Zhang investigated the concept of the hybrid ANN-ARIMA model to obtain precise results as compared to using both models separately [12]. Numerous techniques, which explored the hybrid approach have been used for many years to take advantage of the unique strengths of each of the various types of models. The objective of merging the models is due to the notion that a single model is able to define all the specifics of time series [32]. Mathematically, time-series data can be expressed as a combination of linear and non-linear components [15]:

$$Y_t = Lt + Nt \tag{14}$$

where Y_t shows the time-series data, Lt indicates the linear components, and the non-linear components are represented by Nt.

Mathematically, the neural network model for residual of n input nodes can be expressed as follows:

$$e_t = f(e_{t-1} + e_{t-2}, \dots, e_{t-n}) \tag{15}$$

where f is a non-linear function that is specified by the neural network. With regard to the results of the prediction error of N_t, the combination forecast using the hybrid method can be expressed as:

$$\hat{y}_t = \hat{L}_t \ + \hat{N}_t. \tag{16}$$

N_t is obtained from the predicted values of the ANN model while $\hat{L}_t$ is the forecasted value from ARIMA based on the residual values.

3. Results and Discussion

Kappa Citrine and JMP software were used for the simulation of the models. The experimental results obtained are explained below.

3.1. Results for the Arps Model

Kappa Citrine software was initially used for determining the parameters for the Arps model. The b and D_i values were found to be 1.10 and 0.12 respectively. Subsequently, JMP software was used to construct the prediction model. The second step was to graph a semi-log plot (log q vs. t) to determine the model forecasting equation and parameters. The forecasting equation is given as follows:

$$y = \frac{c}{1 + e^{(-ax^2 - b)}} \qquad (17)$$

where c is the asymptote, a the growth rate while b is the inflection point. The actual and forecasted flow rate values are shown graphically in **Figure 2**.

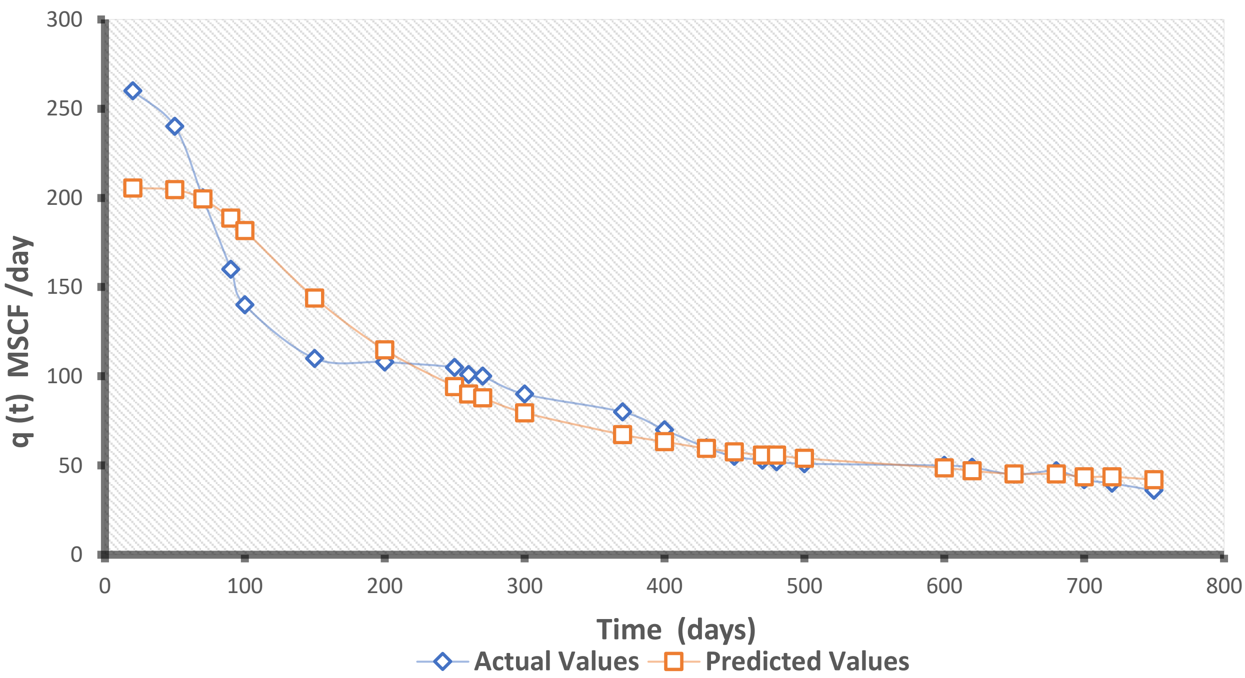


Figure 2: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Arps model

The results for the model appear in some instances to over- and in other instances to underestimate the data. The results concur with literature, which suggests that the weakness of the Arps model is overestimation of results. Tan et al. (2018) [32] in their study highlighted that although the Arps model is simple and fast, it often fails to accurately fit the decline curve of unconventional reservoirs. They further explained that the model often tends to overestimate the EUR for shale gas wells because it assumes that a BDF regime is evident. Paryani et al.

(2018) [3] concurred with these findings, explaining that the drainage area is not constant because the pressure pulse continues to spread from the fracture to other areas of the reservoir volume. Under these conditions, the b value predicted by the Arps model for the actual production data will be greater than 1 as in this case b = 1.10. This in turn leads to inaccurate estimates of reserves.

3.2. Results for the Duong Model

The parameters for the Duong model were q_i = 361.2, a = 1.07 and m = 1.10 respectively. In this instance a log–log linear plot (log q vs. log t) was used. The forecasting equation is given as:

$$y = bx + c \tag{18}$$

where b is the slope and c is the intercept. The actual and forecasted flow rate values can be seen in **Figure 3**.

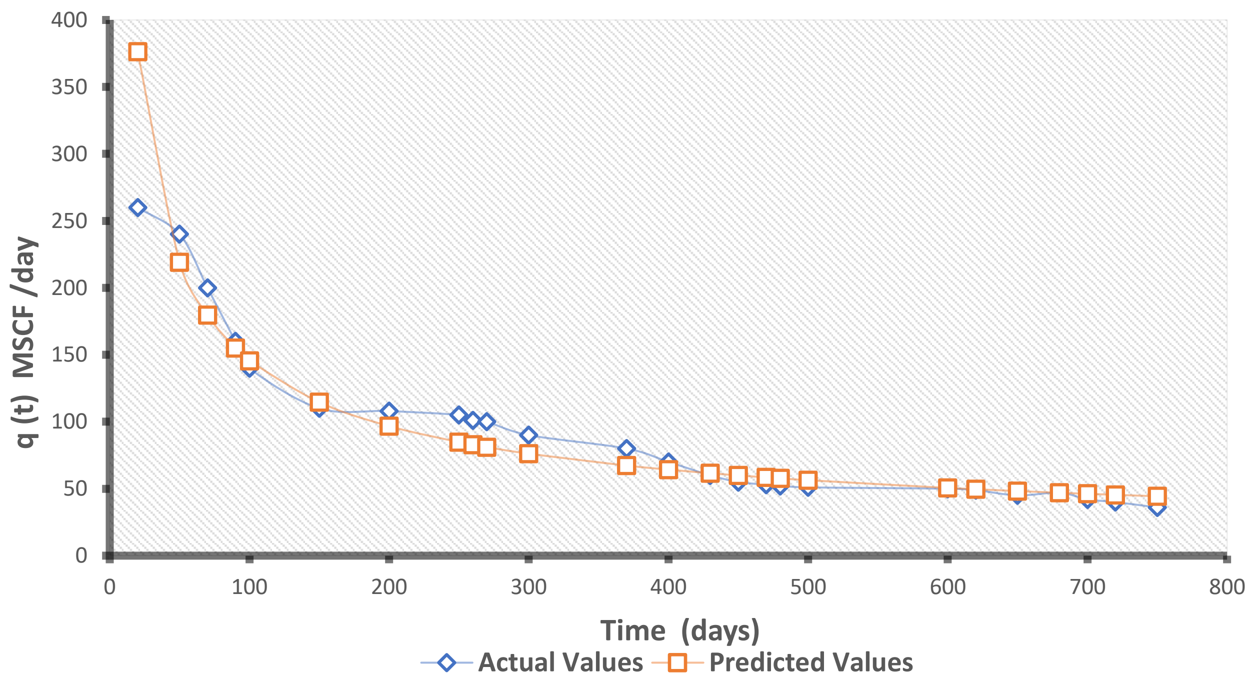


Figure 3: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using Duong's model

The results for the Duong model indicate an overall underestimate of the data. Meyet et al. (2013) [33] mentioned in their work that the Duong model tends to provide the most conservative results. This could also be attributed to the fact that the Duong model tends to be more accurate for linear flows and bilinear–linear flows [28]. Paryani et al. (2018) [3] in their work found that the well fitted with 51% of the historical production data, and that the Duong model fits better with longer and less noisy historical production data.

3.3. Results for the Power Law Exponential (PLE) Model

The parameters used in the model for n and D_i are 0.182 and 0.268 respectively. A log–log plot (log q vs. log t) was used in the model forecasting. The forecasting equation is given as:

$$y = a + be^{cx} \tag{19}$$

where a is the asymptote, b is the scale and c is the growth rate. The actual and forecasted values can be seen in **Figure 4**.

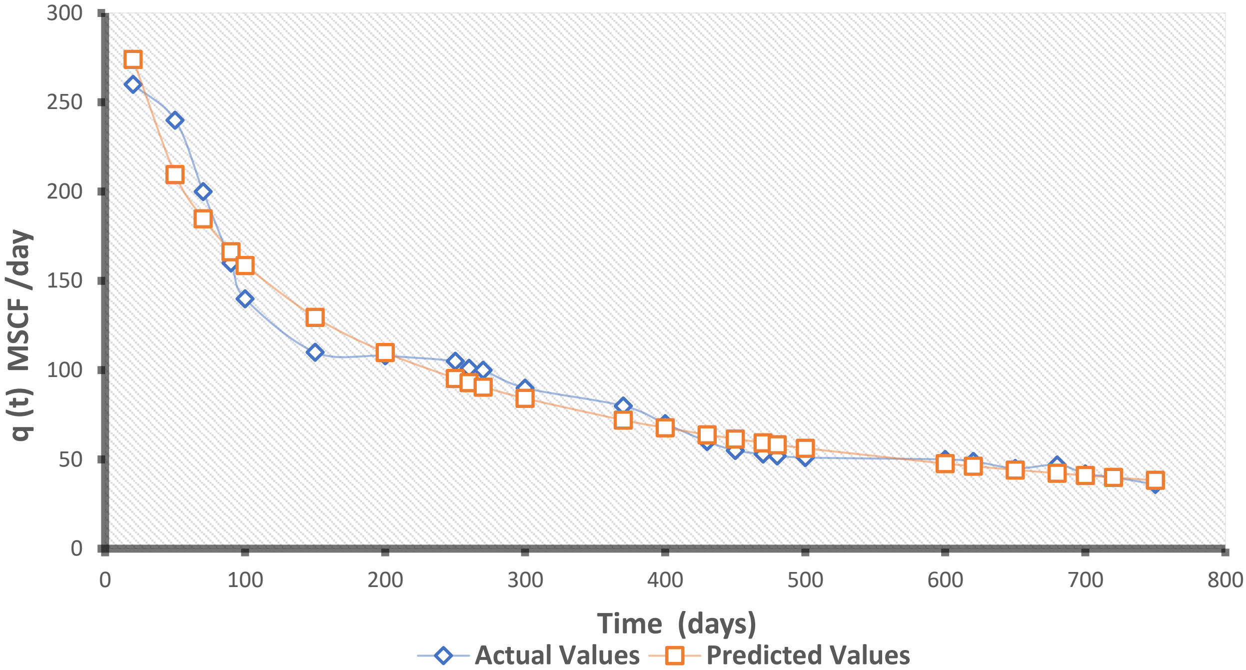


Figure 4: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the PLE model

The results for the PLE model appear to underestimate the data although the PLE considers BDF and TFR, which is an advantage of the model. Furthermore, the model was specifically developed for SGRs, hence it was assumed that the results would be better. This is comparative to the findings by Paryani et al. (2018) [3], as based on their results the PLE consistently gave the lowest forecasts for all the models. It is therefore the most conservative method for production forecasting and reserves estimation. Seshadri and Mattar (2010) [27] concluded that for tight gas wells, the PLE model is complex and non-intuitive. The power law model can result in a non-unique solution due to four degrees of freedom resulting from the four unknown parameters [34].

3.4. Result for the Arps-Duong-PLE Hybrid Model

A plot of $\frac{qt}{qi}$ vs. t was used in the model forecasting. The parameter q_i used was 361.2 which was noted earlier in Duong's model. The forecasting equation is given as:

$$y = a + be^{cx} \tag{20}$$

where a is the asymptote, b is the scale and c is the growth rate. The actual and forecasted values are graphically represented in **Figure 5**.

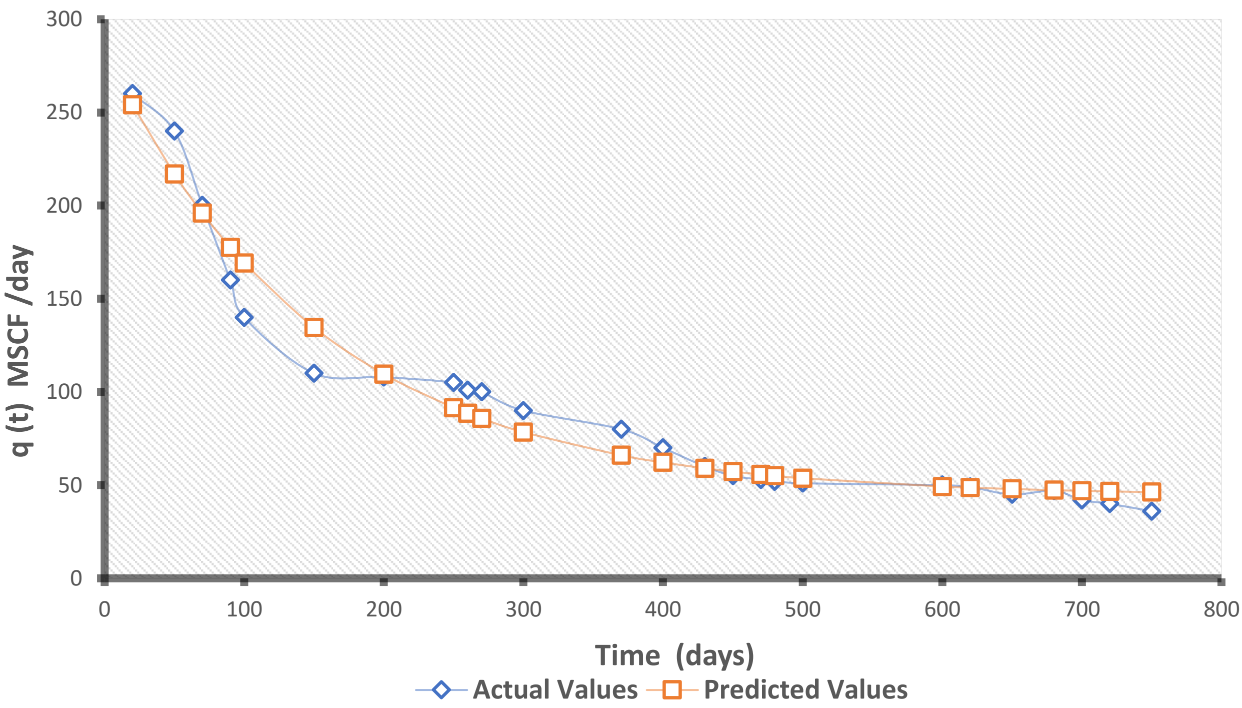


Figure 5: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Arps-Duong-PLE hybrid model

Based on the results, the model appears to over- and underestimate the data. However, the gap between the actual and predicted results is minimised. This could be attributed to both BDF and TFR being considered. In addition, the conservative approach of Duong's and the PLE models along with the inaccurate fitting of the Arps decline curve of unconventional reservoirs could be a contributing factor.

3.5. Result for the Arps-Duong Hybrid Model

A plot of $\frac{qt}{t}$ vs. t was used in the model forecasting. The forecasting equation is given as:

$$y = a(1 - e^{-cx}) \tag{21}$$

where a is the asymptote, b is the scale and c is the growth rate. The actual and forecasted values can be seen in **Figure 6**.

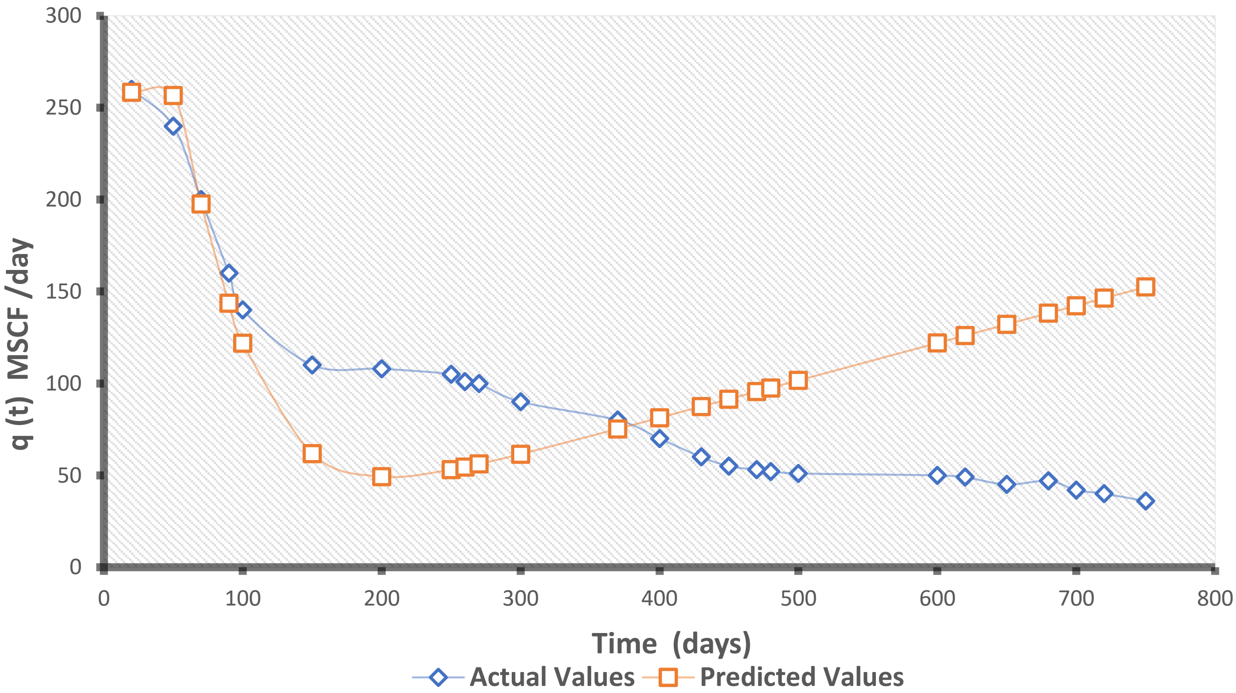


Figure 6: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Arps-Duong hybrid model

The predicted results for the model appear to be severely overestimated from the actual results in the latter stage of production. This would be the result of combining the drawbacks of the two models, which causes the elevated results observed. In line with this, firstly, most shale gas wells rarely reach the boundary-dominated flow regime, hence the Arps model cannot be applied directly to SGRs without significant modifications [32]. Secondly, in the findings of Paryani et al. (2018) [3], extremely high reserves estimates were occasionally observed with the Duong model. The results of Hu et al. (2018) [35] concurred with these results, for the

Austin Chalk wells, whereby the Duong model gave the highest weighted residual of production rate.

3.6. Result for the Arps-Power Law Exponential Hybrid Model

A plot of $\frac{\frac{1}{b}}{qt}$ vs. t was used in the model forecasting. The forecasting equation is given as:

$$y = \frac{c}{1 + e^{(-ax-b)}} \tag{22}$$

where c is the asymptote, b is the inflection point and a is the growth rate. The actual and forecasted values can be seen in **Figure 7**.

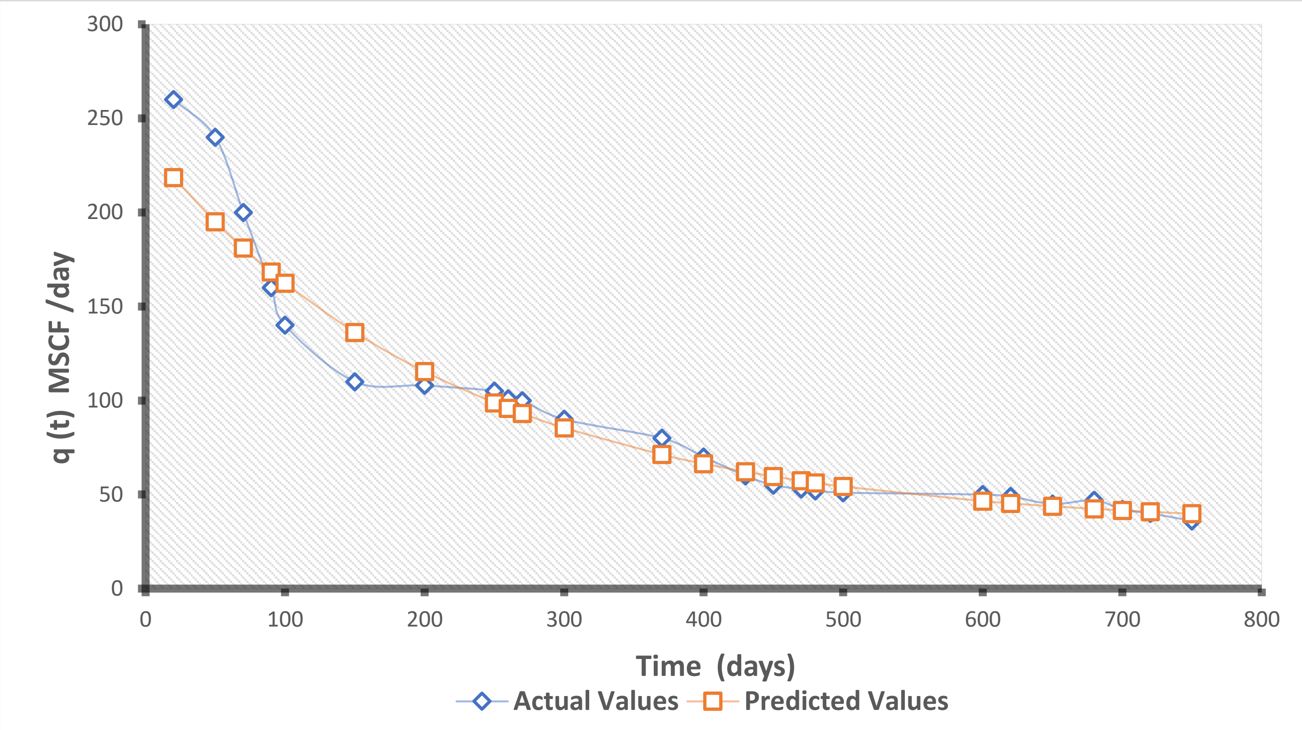


Figure 7: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Arps-PLE hybrid model

The results from the model initially appear to over- and underestimate the data prediction; however, the results tend to move closer to the actual values over time. This would be attributed to the reliability in the Arps model and the fact that the PLE model was developed precisely for SGR. Moreover, both flow regimes are considered and since most shale gas wells rarely reach the boundary-dominated flow regime, the results appear to move closer to the actuals when reaching the TFR. Hence, by combining the models the overestimation of the predicted results is minimised over time.

3.7. The Duong-PLE Hybrid Model

A plot of $\frac{In\, qt}{qm}$ vs. t was used in the model forecasting. The forecasting equation is given as:

$$y = a(1 - be^{-cx}) \tag{23}$$

where a is the asymptote, b is the scale and c is the growth rate. The actual and forecasted values can be seen in **Figure 8**.

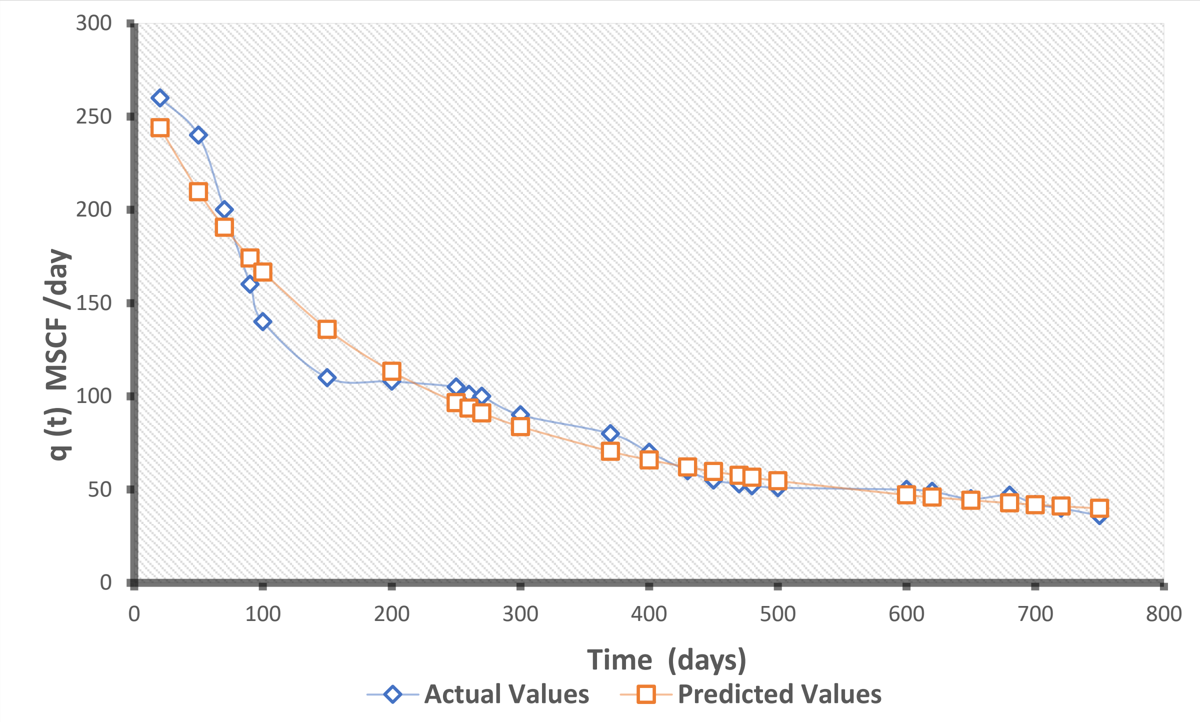


Figure 8: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the Duong-PLE hybrid model

The trend of the results indicate an over- and underestimation. As mentioned by Vanorsdale [36], the PLE and Duong's model will yield an optimistic recovery when the flow regime changes. This trend is clearly evident in the results when combining the models.

3.8. Result for the ARIMA Model

As mentioned earlier under the Research Methodology section, the best fit for the ARIMA model was a (2,1,2), which gave the best forecast values due to having the lowest MSE of 4.82, a low BIC of 8.23 and highest adjusted R^2 of 0,979. The best model is reflected as follows:

$$Y_t = \theta_2 Y_{t-2} + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \varepsilon_t \tag{24}$$

The actual and forecasted values can be seen in **Figure 9**.

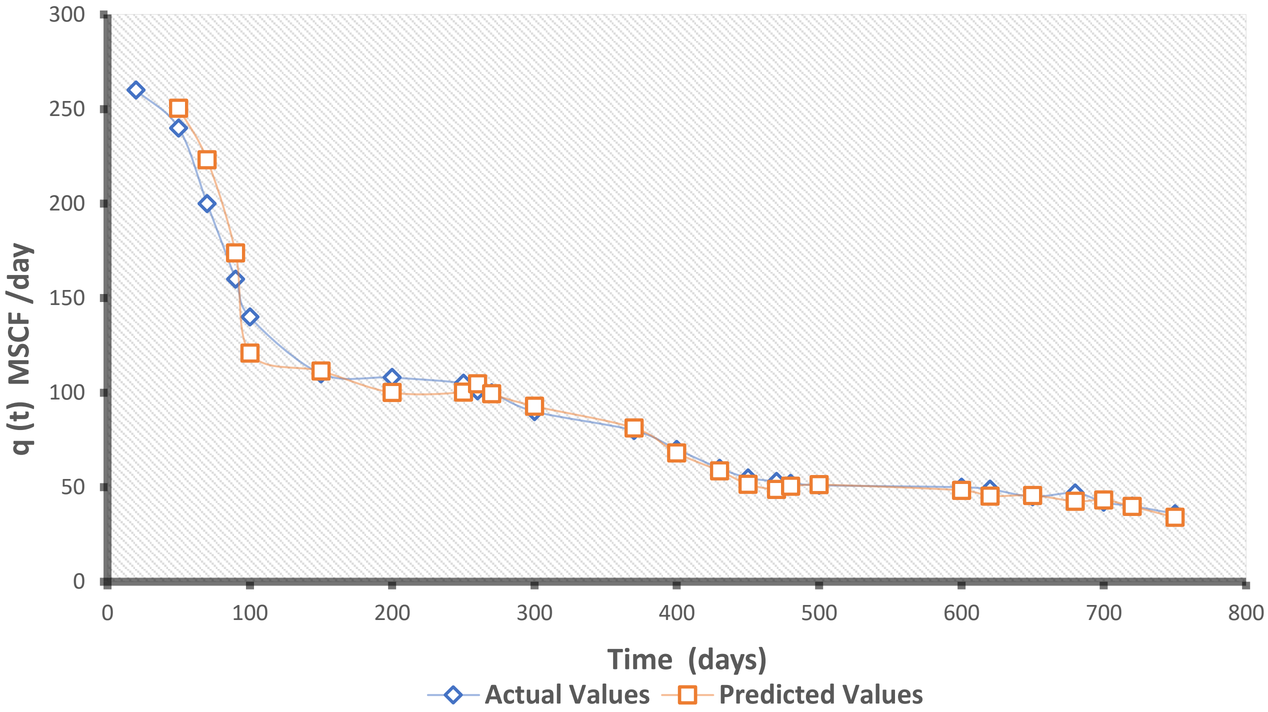


Figure 9: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the ARIMA model

The predicted results from the model appear to follow a close trend to the actual values. Raymond (2007) [37] suggested that ARIMA models have proved to be excellent short-term forecasting models for a wide variety of time series because short-term factors are expected to change slowly. This can explain the reason as to why the ARIMA fared well compared to the other models discussed so far.

3.9. Results for the ANN Model

In the case of this study, a univariate input layer and four nodes gave the best model fit i.e. (1-4-1) for the production flow rate over a period of time. The actual and forecasted values are graphically represented in **Figure 10**.

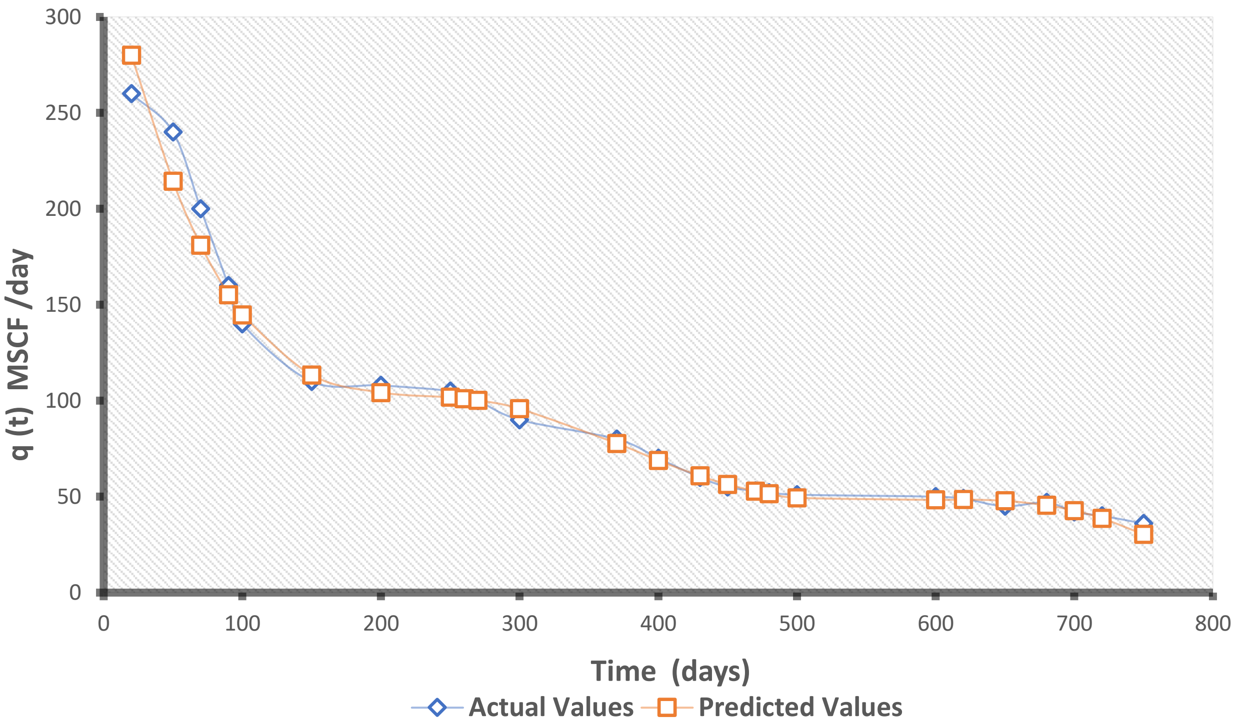


Figure 10: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the ANN model

The predicted results from the model appear to follow a very close trend to the actual values. Zhang (2003) [12] indicated that neural networks are useful for modelling and predicting the properties of time series data. Cybenko (1989) [38] described neural networks as having a universal non-linear function and a relatively good degree of forecasting accuracy. In addition, according to Hill et al. (1996) [39], neural network forecasting provides better results than traditional forecasting methods over monthly as well as quarterly periods.

3.10. Results for the ANN-ARIMA Hybrid Model

The steps employed by Ayub and Jafri (2020) [10] were used to construct the ARIMA-ANN hybrid model. This entailed a two-step process, which involved the following:

In the first step, the ANN is used to predict q_t and residual e_t is produced and provided to the ARIMA to predict the error. In the second step, the predicted q_t by ANN is summed with the error produced by the ARIMA model to give the final predicted values. The equation is as follows:

$$e_t = Y_t - N_t \tag{25}$$

Y_t is time series while N_t is the nonlinear component. ARIMA is used to reproduce e_t to generate the forecast series of q_t. The actual and forecasted values can be seen in **Figure 11**.

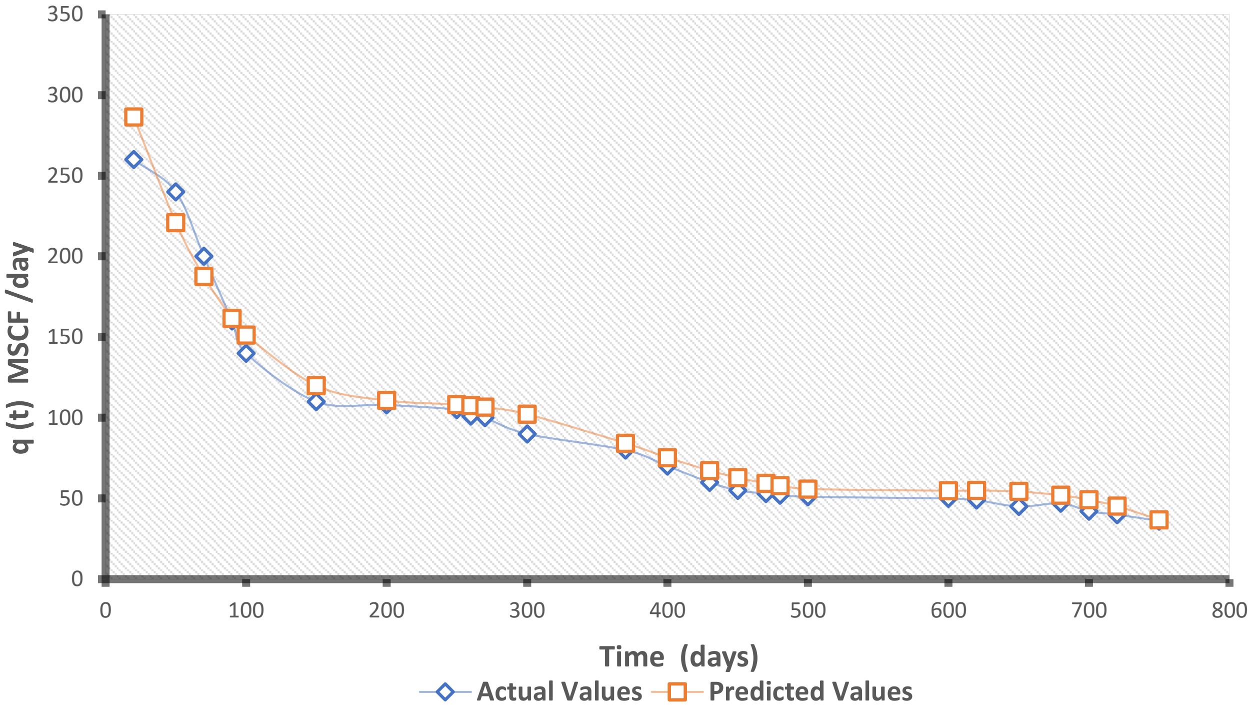
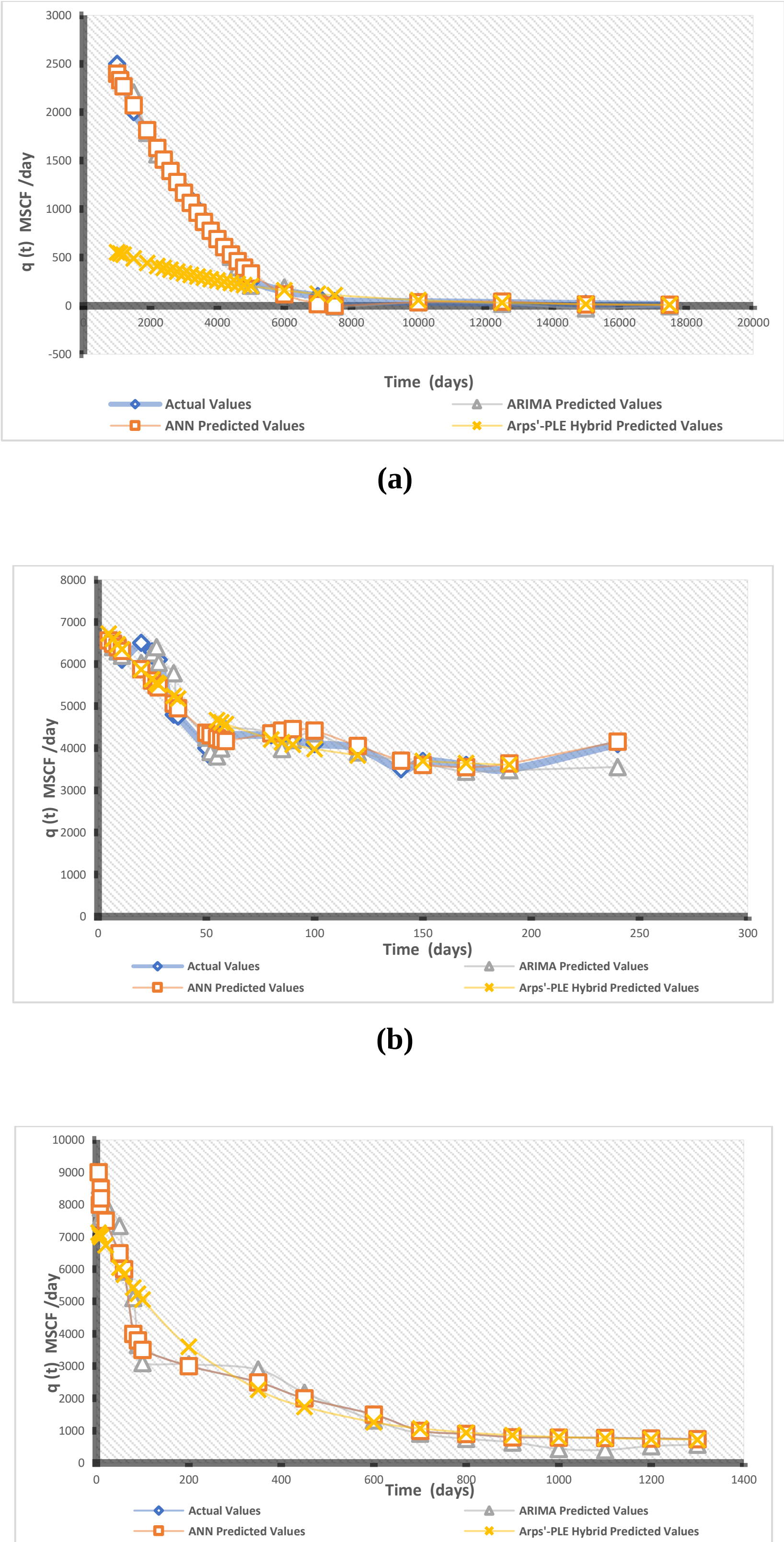


Figure 11: Graphical representation of actual flow rate vs. forecasted flow rate for shale gas production using the hybrid model

The predicted results from the model appear to be overestimated compared to the actual values. This result appears to contradict what has been indicated through the literature. According to Faruk (2010) [40], hybrid methods have a higher degree of accuracy than neural networks. Cybenko (1989) [38] indicated in his work that hybrid models combine the advantages of ARIMA with respect to linear modelling and neural networks in terms of non-linear edge modelling. However, Taskaya-Temizel and Ahmad (2005) [41] made reference in their work that in some circumstances, the single model approach can outperform hybrid models. This has been observed during this study.

3.11. Model Accuracy Evaluation

In order to assess the accuracy of the models, three sets of different production data were used to perform the evaluation. The estimated data was extracted from the work of Adekoya et al. (2009), Brantson et al. (2019) and Tan et al. (2018) [32;42;43]. **Figure 12** illustrates the actual data vs. the predicted data for the ARIMA, ANN and Arps-PLE hybrid models.



(a)

(b)

(c)

Figure 12: Estimated production data to determine accuracy of the different hybrid models (a), (b) and (c) ARIMA vs. ANN vs. Arps-PLE hybrid model [32;42;43]

The results from graphs a, b and c indicate that the ARIMA and ANN models appear to predict the production data very close to the actual values in all three production data; however, this is not the same trend observed for the Arps-PLE hybrid model. The model appears in one instance to underestimate the data and in the other two instances to overestimate the data. Hence, the results prove that with the Arps-PLE hybrid model there is no consistency or accuracy in the prediction of results in the three different production data when compared to the ARIMA and ANN models.

4. Conclusions

The objective of this study was to evaluate the forecasting performance of decline curve hybrid models and ANN-ARIMA hybrid models with Arps', Duong's, PLE decline models, ARIMA and ANN models respectively. The experimental results were obtained using the different prediction models i.e. Arps', Duong's, PLE, Arps-Duong-PLE hybrid, Arps-Duong hybrid, Arps-PLE hybrid, Duong-PLE hybrid, ARIMA, ANN and, lastly, the hybrid ANN-ARIMA model. The following can be concluded from the study:

- The current DCA models, Arps', Duong's and PLE models appear to over- and underestimate the data.
- The DCA hybrid models also did not give the best outcome, which it was assumed they would, in comparison to the individual DCA models. However, the Arps-PLE hybrid model gave the closest predicted results compared to the other DCA hybrid models and the individual models.
- Both the ARIMA and ANN models gave the best predicted results compared to all the models evaluated in this study. However, when both models were combined into the ANN-ARIMA hybrid model the strengths of both models referenced in literature did not provide accurate predictive data. The result was an overestimation in the production flow rate.
- Overall, the models which gave predicted values closest to the actuals in order of rank were the ARIMA, ANN and the Arps-PLE hybrid model.
- In the model accuracy evaluation, the Arps-PLE hybrid model did not provide a consistent prediction. The model under- and overestimated the production data compared to the ARIMA and ANN models.

In conclusion, this study contradicted the findings from literature which indicated that hybrid models have a higher degree of accuracy. However, the study concurred with Taskaya-Temizel and Ahmad (2005) [41], whereby in certain circumstances the single model approach can outperform the hybrid models. Future investigation should therefore validate the ARIMA and ANN models for SGR decline forecasting using the factors R^2, MSE and RMSE.

References

1. Fetkovich, M. J. 1980. Decline curve analysis using type curves. *J. Pet. Technol. 32*, 1065–1077.
2. Tan, L., Zuo, L. and Wang, B. 2018. Methods of decline curve analysis for shale gas reservoirs. *Energies 11*, 552.
3. Paryani, M., Ahmadi, M., Awoleke, O. and Hanks, C. 2018. Decline Curve Analysis: A Comparative Study of Proposed Models Using Improved Residual Functions. *J. Pet. Environ. Biotechnol*. *9*, 362.
4. Ariyo, A. A., Adewumi, A. O. and Ayo, C. K. 2014. Stock price prediction using the ARIMA model. In 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation. *IEEE* March, 106–112.
5. Babu, C. N. and Reddy, B. E. 2014. A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. *Applied Soft Computing 23*, 27–38.
6. Guresen, E., Kayakutlu, G. and Daim, T. U. 2011. Using artificial neural network models in stock market index prediction. *Expert Systems with Applications 38*(8), 10389–10397.
7. Bacaër, N. 2011. Verhulst and the logistic equation (1838). *In A Short History of Mathematical Population Dynamics*; Springer: London, UK, 35–39.
8. Taskaya-Temizel, T. and Ahmad, K. 2005. Are ARIMA neural network hybrids better than single models? In *Proceedings of the International Joint Conference on Neural Networks* (IJCNN 2005), Montr’eal, QC, Canada, 31 July–4 August 2005.
9. Faruk, D. Ö. 2010. A Hybrid Neural Network and ARIMA Model for Water Quality Time Series Prediction. *Eng. Appl. Artif. Intell*. *23*, 586–594.
10. Ayub, S. and Jafri, Y. Z. 2020. Comparative Study of an ANN-ARIMA Hybrid Model for Predicting Karachi Stock Price. *American Journal of Mathematics and Statistics 10*(1), 1–9.
11. Contreras, J., Espinola, R., Nogales, F. J. and Conejo, A. J. 2003. ARIMA models to predict next-day electricity prices. *IEEE transactions on power systems 18*(3), 1014–1020.
12. Zhang, G. P. 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing 50*, 159–175.

13. Fattah, J., Ezzine, L., Aman, Z., El Moussami, H. and Lachhab, A. 2018. Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management 10*, 1–9.
14. Shamsuddin, S. M., Sallehuddin, R. and Yusof, N. M. 2008. Artificial neural network time series modelling for revenue forecasting. *Chiang Mai J. Sci*. *35*(3), 411–426.
15. Kaastra, I. and Boyd, M. 1996. Designing a neural network for forecasting financial. *Neurocomputing 10*, 215–236.
16. Dhini, A., Riefqi, M. and Puspasari, M. A. 2015. Forecasting analysis of consumer goods demand using neural networks and ARIMA. *International. Journal of Technology 6*, 872–880.
17. Bacha, H. and Meyer, W. 1992. A neural network architecture for load forecasting, *IEEE/INNS International Joint Conference on Neural Networks 2,* 442–447.
18. Boah, E. A., Borsah, A. A. and Brantson, E. T. 2018. Decline Curve Analysis and Production Forecast Studies for Oil Well Performance Prediction: A Case Study of Reservoir X. *Int. J. Eng. Sci. (IJES) 7*, 56–67.
19. Ali, T. A. and Sheng, J. J. 2015. Production decline models: A comparison study. In Proceedings of the SPE Eastern Regional Meeting, Morgantown, WV, USA, 13–15 October 2015.
20. Yuhu, B., Guihua, C., Bingxiang, X., Ruyong, F. and Ling, C. 2016. Comparison of typical curve models for shale gas production decline prediction. *China Pet. Explor*. *21*, 96–102.
21. Li, P., Hao, M., Hu, J., Ru, Z. and Li, Z. 2018. A new production decline model for horizontal wells in low-permeability reservoirs. *J. Pet. Sci. Eng. 171*, 340–352.
22. Qu, Z. and Lin, J. E. (Eds.) 2018. *Proceedings of the International Field Exploration and Development Conference 2017,* Springer: Singapore.
23. Duong, A. N. 2011. Rate-decline analysis for fracture-dominated shale reservoirs. *SPE Reserv. Eval. Eng*. *14*, 377–387.
24. Lee, K. S. and Kim, T. H. 2016. *Integrative Understanding of Shale Gas Reservoirs*, Springer: Heidelberg, Germany.
25. Ilk, D., Rushing, J. A., Perego, A. D. and Blasingame, T. A. 2008. Exponential vs hyperbolic decline in tight gas sands: Understanding the origin and implications for reserve estimates using Arps decline curves. In Proceedings of the SPE Annual Technical Conference and Exhibition, Denver, CO, USA, 21–24 September 2008.

26. McNeil, R., Jeje, O. and Renaud, A. 2009. Application of the power law loss-ratio method of decline analysis. In Proceedings of the Canadian International Petroleum Conference, Calgary, AB, Canada, 16–18 June 2009.
27. Seshadri, J. N. and Mattar, L. 2010. Comparison of power law and modified hyperbolic decline methods. In Proceedings of the Canadian Unconventional Resources and International Petroleum Conference, Calgary, AB, Canada, 19–21 October 2010.
28. Kanfar, M. S. and Wattenbarger, R. A. 2012. Comparison of Empirical Decline Curve Methods for Shale Wells. In Proceedings of the SPE Canadian Unconventional Resources Conferences, Calgary, AB, Canada, 30 October–1 November 2012.
29. Li, P., Hao, M., Hu, J., Ru, Z. and Li, Z. 2018. A new production decline model for horizontal wells in low-permeability reservoirs. *J. Pet. Sci*. *Eng*. *171*, 340–352.
30. Büyükşahin, Ü. Ç. and Ertekin, Ş. 2019. Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. *Neurocomputing 361*, 151–163.
31. Khashei, M., Bijari, M. and Ardali, G. A. R. 2009. Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs). *Neurocomputing*, *72*(4–6), 956–967.
32. Tan, L., Zuo, L. and Wang, B. 2018. Methods of decline curve analysis for shale gas reservoirs. *Energies*, *11*(3), 552.
33. Meyet Me Ndong, M.P., Dutta, R. and Burns, C. 2013. Comparison of decline curve analysis methods with analytical models in unconventional plays. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 30 September–2 October 2013.
34. Ali, T. A. and Sheng, J. J. 2015. Production decline models: A comparison study. In Proceedings of the SPE Eastern Regional Meeting, Morgantown, WV, USA, 13–15 October 2015.
35. Hu, Y., Weijermars, R., Zuo, L. and Yu, W. 2018. Benchmarking EUR estimates for hydraulically fractured wells with and without fracture hits using various DCA methods. J. Pet. Sci. Eng. *162*, 617–632.
36. Vanorsdale, C. R. 2013. Production decline analysis lessons from classic shale gas wells. In Proceedings of the SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 30 September–2 October 2013.
37. Raymond, Y. C. 1997. An application of the ARIMA model to real-estate prices in Hong Kong. *Journal of Property Finance 8*(2), 152–163.

38. Cybenko, G. 1989. Approximation by Superposition's of a Sigmoidal Function. Mathematics of Control, *Signals and Systems*, *2*(4), 303–314.
39. Hill, T., O'Connor, M. and Remus, W. 1996. Neural Network Models for Time Series Forecasts. *Management Science*, *42*(7), 1082–1092.
40. Faruk, D. Ö. 2010. A Hybrid Neural Network and ARIMA Model for Water Quality Time Series Prediction. *Engineering Applications of Artificial Intelligence*, *23*(4), 586–594.
41. Taskaya-Temizel, T. and Ahmad, K. 2005. Are ARIMA Neural Network Hybrids Better than Single Models? In: *Proceedings of the IEEE International Joint Conference on Neural Networks 5*, 3192–3197.
42. Adekoya, F. 2009. Production Decline Analysis of Horizontal Well in Gas Shale Reservoirs. Master's Thesis, West Virginia University, Morgantown, WV, USA.
43. Brantson, E. T., Ju, B., Ziggah, Y. Y., Akwensi, P. H., Sun, Y., Wu, D. and Addo, B. J. 2019. Forecasting of Horizontal Gas Well Production Decline in Unconventional Reservoirs using Productivity, Soft Computing and Swarm Intelligence Models. *Nat. Resour. Res. 28*, 717–756.

Accuracy Assessment of Single and Hybrid Models for Predicting Shale Gas Production

Prinisha Manda*† and Diakanua Nkazi†

† Oil and Gas Production and Processing Research Unit, School of Chemical and Metallurgical Engineering, University of the Witwatersrand, Johannesburg 2000; South Africa

Abstract

The accuracy of a predictive tool determines the levels of trust in the model and its attraction for commercial usage. The study examined the single and hybrid model approach for shale gas production. Multilayer perception ANN, ARIMA and Arps-Power Law Exponential hybrid decline models were developed to predict shale gas production and were compared with the already developed Arps Decline and Power Law Exponential Decline (PLE) models. Using a trial and error approach, a MLP network with 4 neurons in the hidden layer was attained in the ANN structure to predict shale gas production; Whereas for the ARIMA model, the number of nodes which showed the best performance indicated a (2,1,2) for the two sets of data, respectively. Evaluation of the RMSE values for the models showed that the Arps-Power Law Exponential hybrid decline model had the lower percentage error in conjunction with good accuracy. The study found the Arps-Power Law Exponential hybrid decline model to be a good forecaster of shale gas production and that hybrid models deliver better accuracy over single models. Future revision of model assumptions may improve its accuracy, and make the Arps-Power Law Exponential hybrid decline model an attractive predictive tool.

1. Introduction

Shale gas reservoirs (SGRs) have developed into a vital source for providing natural gas with; hydraulic fracking being the process used in the extraction of the shale gas.

1 However, consistently forecasting the reserves of shale basins over a long period has been a challenge.

2 The gas industry hence wants simple, beneficial and quick means of forecasting and assessing production data from reserves. Therefore, Decline Curve Analysis (DCA) has been a smart alternative to other methods.[2]

An accuracy assessment showed that decline curve modelling influences the estimated ultimate recovery (EUR) of SGRs and that the different decline curve models approximately differ from EUR results which are either over or under estimated.

3 Previous research has shown that production time considerably influences the EUR which is dependent on the decline curve model being used.

4 It is apparent from the assessment that there are benefits in using existing DCA models. However, they also have constraints associated with them, which have to be addressed.[3]

A proposed hybrid model approach was examined since some studies proved it to produce higher accuracy. However, based on the findings from the study on production decline prediction using hybrid models, the results showed that ANN, ARIMA and Arps-Power Law Exponential (Arps-PLE) hybrid rate decline models showed better accuracy.

5 The finding concurred with Taskaya-Temizel et al. 6 who showed that under certain conditions, single model methodology can outperform hybrid models. This paper will focus on the accuracy and validation of the single models: Arps decline, PLE decline, ANN, ARIMA, and the hybrid models: Arps-PLE hybrid decline and ANN-ARIMA hybrid for shale gas production. The main ideas are that a researcher should; a) use the goodness-of-fit statistical assessment to evaluate the accuracy of the models, b) validate the models using the MAPE values and c) summarise findings from the research study.

2. Models Overview

2.1. Arps Decline Curve Model

The Arps decline curve model analysis is the most repeatedly used technique of estimating ultimate recoverable reserves and future performance.

7 Paryani et al. details this to show a consistent history match (even with b > 1) as well as its straightforwardness. The model development was founded on certain focal assumptions: that historical operating conditions will stay unaffected, a well is produced at or near capacity, and that the well's drainage remains constant and is produced at a constant bottom-hole pressure.[9] Particularly, the Arps model is only valid in pseudo-steady flows when the flow regime moves from linear flows to boundary-dominated flows (BDF).[10] This shows that Arps Equations are not applicable to the production forecasting of the total decline process of horizontal wells in low-permeability reservoirs.[11] The Arps decline curve analysis can be categorised into three types: exponential Equation (1), hyperbolic Equation (2), and harmonic Equation (3).[12,13]

$$q = q_i e^{-Dt} \tag{1}$$

$$q = \frac{q_i}{(1 + bD_i t)^{\frac{1}{b}}} \tag{2}$$

$$q = \frac{q_i}{1 + D_i} \tag{3}$$

where q is the flow rate in STB/day or Mscf/day, q_i is the initial flow rate in STB/day or Mscf/day, D is the decline constant, D_i is the initial decline constant (both measured in days $^{-1}$), and b is the decline exponent.[12,13]

2.2. Power Law Exponential Model (PLE)

Ilk et al.[14] introduced the PLE model, which is an addition of the exponential Arps equation for the decline degree in SGRs. The PLE model was developed specifically for SGRs and estimates the degree of decline with a power law decline. It considers production data in both transient and boundary-dominated regions, without being oversensitive to the remaining reserve estimates.[15] Tan et al.[15] highlighted that this model can model transient radial and linear flows, while Kanfar et al.[16] showed that the PLE model is reliable for linear flow and bilinear flow, followed by linear flow, and linear flow, followed by BDF, or bilinear flow, followed by linear flow and finished with BDF flow. Vanorsdale[17] inferred that when the flow regime changes throughout the initial ten years of the well, the model would yield a very optimistic recovery. The model characterizes the decline rate by infinite time, D_∞ which is defined as a "loss ratio" (which is assumed to be constant from Arp).[11] The production rate is derived as follows:

$$\frac{q}{dq/dt} = -b \tag{4}$$

$$b = D_\infty D_i t^{-(1-\hat{n})} \tag{5}$$

where dq/dt is the slope, D_i is the decline rate over a long-term period, and $\hat{n}$ is the time exponent. By substituting the above equations, the production rate is attained:

$$q(t) = \hat{q}_i e^{[-D_\infty t - \hat{D}_i t^{\hat{n}}]} \tag{6}$$

In the above equation, there are four unknown variables: $\hat{q}_i, \hat{D}_{i,}\ D_\infty\ and\ \hat{n}$, which adds to the number of degrees of freedom and possibly will be difficult to use or solve.[18] Johnson et al.[19] referenced in their work that the D_i parameter is hard to solve. However, there are benefits to using this model; in that the additional variables allow for both transient and boundary flow, and the equation for production rate seems comparable to the Arps exponential equation.[8]

2.3. The ARIMA Model

The Autoregressive Integrated Moving Average (ARIMA) methodology for the study of time series analysis was developed by Box and Jenkins.[20] The Box–Jenkins method encompasses three repetitive steps; being; model identification, parameter estimation and diagnostic checking.[21] The three-step model construction process is usually repeated several times until an acceptable model is finally selected and can then be used for forecasting purposes.[21]

For the ARIMA model, the future value of a variable is presumed to be a linear function of several past observations and random errors.[21] During past years, academics have been focused more on linear models, due to their ease of comprehension and application.[22] An ARIMA model is labelled using three terms (p, d, q), where:[22]

7. p, number of autoregressive terms
8. d, number of differences, and,
9. q, number of moving averages.

2.3.1. The Autoregressive Process

This process assumes Y_t is a linear function of the previous values and is given by the Equation (7).[22]

$$Y_t = \propto_1 Y_{t-1} + \varepsilon_t \tag{7}$$

Generally, each observation comprises a random component i.e. a random shock, ε and a linear combination of the previous observations. $\propto_1$ in the equation is the self-regression coefficient.[22]

2.3.2. The Integrated Process

The integrated processes are the model of non-stationary series. A distinction of order 1 assumes that the difference between two succeeding values of Y is constant. An integrated process is defined by Equation (8).[22]

$$Y_t = Y_{t-1} + \varepsilon_t \tag{8}$$

where ε_t is a white noise.[22]

2.3.3. The Moving Average Process

The moving averaging method is a linear grouping of the current disturbance with one or more previous perturbations. The moving average order designates the number of previous periods embedded in the existing value. Consequently, the moving average is defined by Equation (9).[22]

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{9}$$

2.4. ANN Model

ANN was initially developed to simulate the basic biological neural system i.e. the human brain that comprises a number of interconnected neutrons or nodes.[23] Each node accepts an input signal, which is the total "information" from other nodes, processes it locally through an activation or transfer function and produces a transformed output signal to other nodes.[23] According to Lari[23,] while each single neuron implements its function rather slowly and imperfectly, jointly, a network can complete an amazing number of tasks efficiently.[23] This processing function makes ANN a dominant computational tool that is capable of learning from examples and is then able to simplify it with patterns encountered previously.[23]

2.4.1. ANN Architecture

There have been many different ANN models recommended from the 1980s. However, the most used models are the multilayer perceptions (MLP), the Hopfield networks and Kohonen's self-organising networks.[24] For this study, the MPL will be used because it can be used in a wide range of situations, particularly in forecasting, because of its essential ability in random input-output mapping.[24] The MLP model comprises three interconnected layers, the input layer, the hidden layer, and the output layer.[24] The pattern of the recommended network with Z neurons in the hidden layer is presented in **Figure 1**.

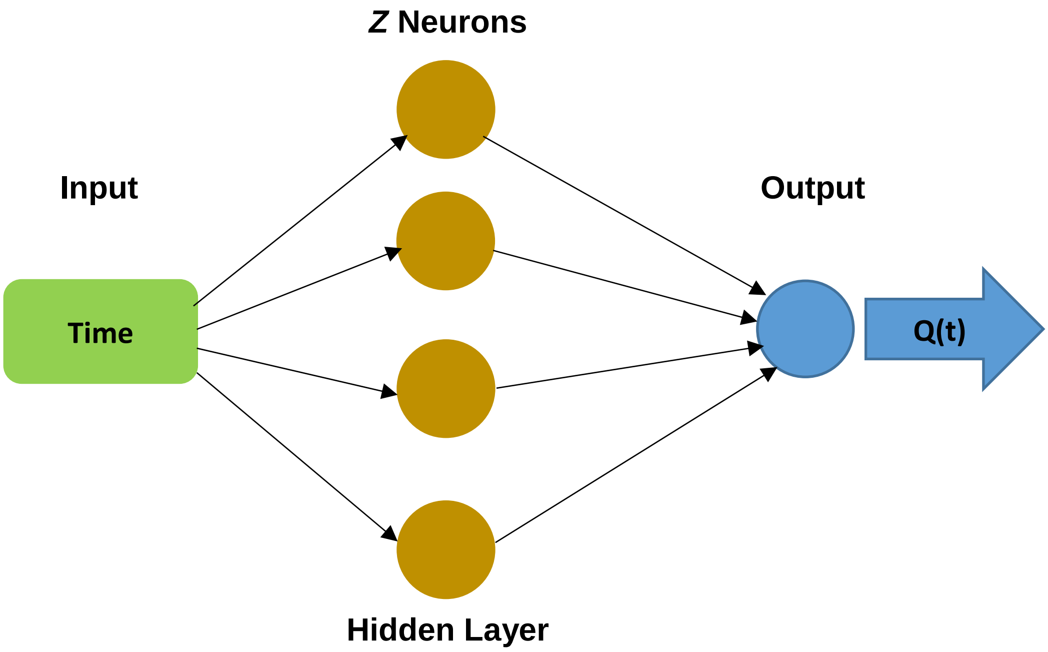


Figure 1: Graphic illustration of the ANN (1- Z -1) model applied in this study

The fundamental unit of any ANN is the neuron or node, which is the processor. Each neuron/node is capable of summing many inputs x1, x2 ... x3 whether these inputs are from a database or from other nodes, with each input adjusted by an adjustable connection weight.[25]

The connection that occurs in the output and input layers follow Equation (10).[25]

$$Y_t = \propto_0 + \sum_{j-1}^{q} \propto_j g\left(\beta_0 j + \sum_{I=1}^{p} \beta_i j Y_t - i\right) + \varepsilon_t \quad \textbf{(10)}$$

where $\propto_j$ (*j* = 1,2,3, ..., q) and $\beta_i j$ (*i* = 1,2,3, ..., p; *j* = 1,2,3, ..., q) are the parameters of the model, *p* is the number of input nodes, and q is the number of hidden nodes.[25]

2.4.2. Selection of the Best ANN Network Architecture

The best selection should produce small errors for training and tested data. Therefore, the selection of the amount of input and hidden neurons/nodes becomes critical.[25] The topology of an MLP network related with the number of neurons/nodes in the hidden layer has a substantial effect on prediction accuracy and the generalization ability of the network, and therefore should be improved.[26] However, there is no common rule for the determination of the optimal topology of an MLP network and it is usually determined through the trial-and-error method.[27]

2.5. Hybrid Models

As mentioned earlier, the hybrid model philosophy is a new method in predicting a decline in shale gas production. There is minimal literature explaining the hybrid model method in combining decline curve models used in the prediction process. In this article a proposed hybrid decline model which incorporates the Arps and PLE models is evaluated. The Arps model considers linear to BDF, while the PLE model considers TFR flows. As mentioned by Kuila et al.[28] gas flow in SGRs is defined by a combination of mechanisms acting at different scales. These are: [28]

- Desorption from kerogen and clay surfaces, and subsequent surface diffusion of the adsorbed gas molecules under a pressure gradient.
- Knudsen diffusion and slip flow in micropores, and
- Darcy flow in larger meso-and macropores.

According to Huang et al.[29] principal analysis shows that the permeability of shale is in a range of 1×10^{-9}–1×10^{-3} μm^2 and flow in extremely low permeability shales undergoes a transition of flow regimes owing to the significant effect of collisions between molecules and pore walls on gas transport.[29]
Gas flow regimes can be classified into four groups, depending on the Knudsen number (Table 1).[29]

Table 1: Gas flow regime classification founded on the Knudsen number

Knudsen Number	**Kn ≤ 0.001**	**0.001 < Kn ≤ 0.1**	**0.1 < Kn ≤ 10**	**Kn > 10**
Flow regime	Continuum flow	Slip Flow	Transition flow	Free-molecule flow

In continuum flow regime, no-slip boundary condition is effective and gas flow is linear. As Kn increases, the rarefaction effects become more distinct and the continuum assumption breaks down eventually.[29] As the Knudsen number increases, either by increasing the mean free path of gas (i.e., gas flowing at very low pressure) or by decreasing the pore size, the contribution of the Knudsen layer increases. It goes to a transition flow and then finally to pure Knudsen or free molecule flow.[28] Considering the different flow regime changes which occur during the shale extraction process, the hybrid decline Arps-PLE model considers flows in the linear, transition and free molecule regions. The proposed hybrid decline model equation is given as:

$$t[-D_\infty - D_i\hat{n}] = \frac{-\frac{1}{b}}{q_t} In(1 + bD_i) \tag{11}$$

where q_i is the flow rate in STB/day or Mscf/day, t is the time in days, D_∞ the decline rate over a long-term period, D_i the initial decline constant (both measured in days^{-1}), $\hat{n}$ is the time exponent and b is the decline exponent.

The hybrid model concept has commonly been used in the finance sector, specifically for stock price forecasting, which involves combining the ANN and ARIMA models. According to Faruk[30] hybrid methodologies tend to have a greater degree of accuracy than neural networks. The ARIMA model can distinguish patterns of time-series well;

but not data patterns which are non-linear; whereas ANN can only handle non-linear data. Therefore, hybrid models link benefits of ARIMA with respect to linear modelling and ANN, in terms of non-linear edge modelling.[3] Under certain conditions, the single model approach can outperform hybrid models.[6]

The equation for time-series data can be expressed as a combination of linear and non-linear components, as follows:[31]

$$Y_t = L_t + N_t \tag{12}$$

where Y_t shows the time-series data, L_t designates the linear aspects, and the non-linear aspects are denoted by N_t.[31]
The neural network (NN) model for residual of n input nodes can be expressed as follows:[31]

$$e_t = f(e_{t-1} + e_{t-2}, \dots, e_{t-n}) \tag{13}$$

where f is a non-linear function that is identified by the NN. With regard to the results of the prediction error of N_t, the combination forecast using the hybrid method can be expressed as:[31]

$$\hat{y}_t = \hat{L}_t + \hat{N}_t. \tag{14}$$

According to Taskaya-Temizel et al.[6] two factors prevent the hybrid ANN-ARIMA method from providing good results. Firstly, the assumption of the existence of a relationship between the components of the linear and non-linear components in the data can cause performance degradation, as other model relationships (e.g., multiplicative) may exist within the data instead of linear/non-linear relationships. Secondly, no one can assure that the residual of the linear components will have effective non-linear patterns. Their findings showed that hybrids are not always superior and therefore the model selection process remains a vital step despite the popularity of hybrid models.[6]

Dhini et al.[32] concurred that the hybrid method does not always give better results than the single methods, as the neural network method often outperformed the hybrid method. Some of the possible causes for this are the basic assumptions used in the method, as well as the possibility that the residual from the linear components may not be non-linear.[33] Chao et al.[33] found that the hybrid model is capable of outperforming each component model used in isolation. Granger[34] highlighted that for a hybrid model to produce higher forecasts, the component model should be suboptimal.

3. Methodology

The methodology for the research study followed a three step process, **Figure 2** summarises the steps and key findings. The initial stage involved an evaluation and sensitivity analysis of current decline curve models used to predict the decline of shale gas production. The assessment found benefits in using the existing decline curve analysis (DCA) models. However, they too have restrictions linked with them which have to be addressed. A new proposed hybrid model approach was examined and from previous studies conducted has proved to provide higher accuracy, but this approach has not been evaluated for SGR.

The second step was to develop and evaluate the hybrid model approach for decline curve models. The development process involved combining the Arps, Power Law Exponential and Duong's decline models (hybrid models) which were chosen based on the evaluation process in the initial stage. The models was then compared to ANN-ARIMA hybrid model. The findings indicated that DCA hybrid models do not give the best outcome, which it was assumed they would. However, the Arps-PLE hybrid decline model gave the closest predicted results compared to the other DCA hybrid models. Also from the study, the ANN and ARIMA models gave the best predicted results compared to all the models evaluated. However, when both models were combined into the ANN-ARIMA hybrid model, the strengths of both models referenced in the literature did not provide accurate predictive data.

The third stage of the study will involve confirming the findings from the second step of the study by assessing the accuracy and validity of the ANN, ARIMA, Arp's, PLE, Arps-PLE hybrid decline and ANN-ARIMA hybrid models. This will involve using the R^2, RMSE and MAPE parameters.

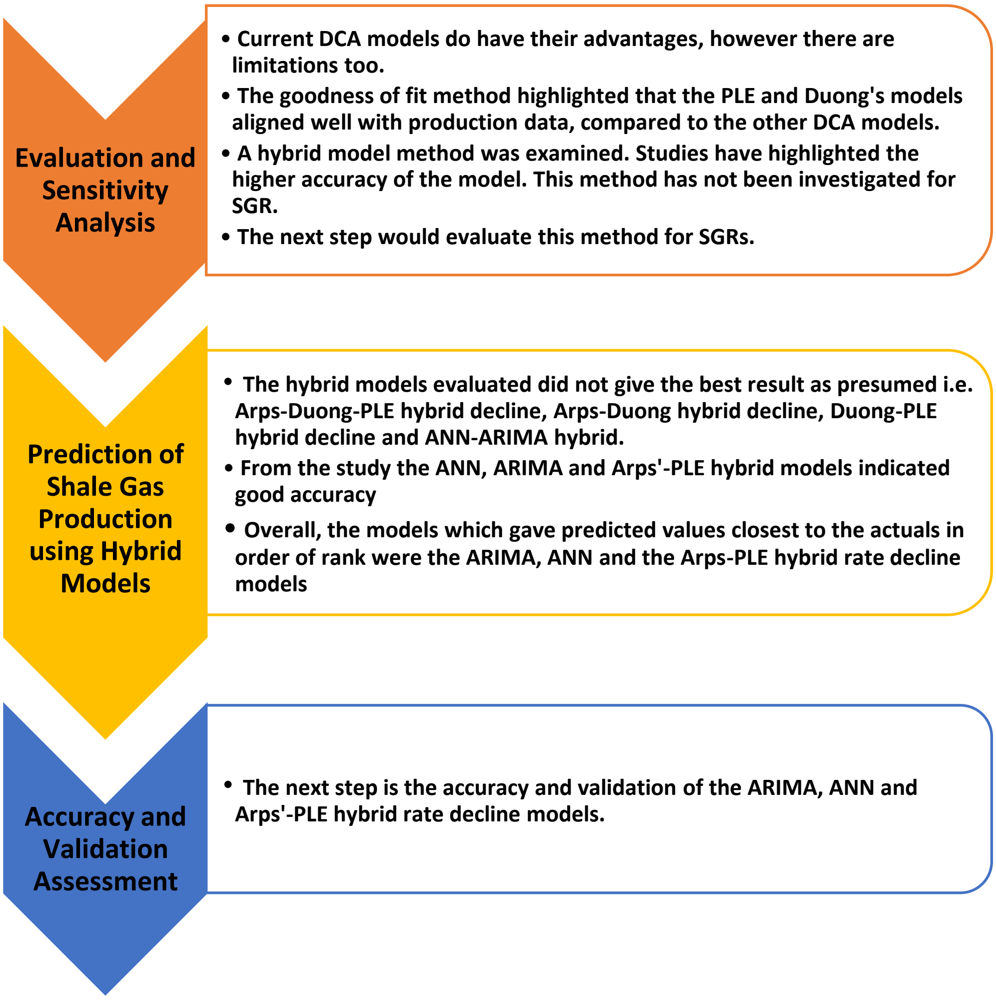


Figure 2: Methodology for optimal model validation

4. Results and Discussion

In order to determine the accuracy and validate the results obtained for the forecasting models, Root Mean Square Error (RMSE), Mean Standard Error (MSE), Mean Absolute Percentage Error (MAPE) and Correlation Coefficient (R^2) between the actual and predicted values. The RMSE, MSE and MAPE and R^2 are defined as follows:

$$RSME = \sqrt{\frac{\sum_{t=1}^{n}(Pred_t - A_t)}{n}} \quad \textbf{(15)}$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(A_t - Pred_t)^2 \quad \textbf{(16)}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{A_t - Pred_t}{A_t} \quad \textbf{(17)}$$

$$R^2 = \left(\frac{\sum_{t=1}^{n}(Pred_t - \overline{Pred_t})(A_t - \overline{A_t})}{\sum_{t=1}^{n}(Pred_t - \overline{Pred_t})^2 \sum_{t=1}^{n}(A_t - \overline{A_t})^2}\right) \quad \textbf{(18)}$$

where $Pred_t$ is the predicted value and A_t is the actual value. $\bar{A}_t$ and $\overline{Pred}_t$ are the average of the actual and predicted values respectively.

In the study, different networks (ANN) and number of orders (ARIMA) were formulated and applied to shale gas production. The number of neurons/nodes in the hidden layer varied from 1 to 15 and each network was repeatedly run 20 times. In the case of the ARIMA model, the number of nodes were varied and the best fit evaluated.

Table 1 summaries the performance results for the ANN model for the different numbers of neurons/nodes in the hidden layer for the Canon Shale Well and Marcellus Shale Well and **Table 2** summarises the results for the ARIMA model. The developed Arps-PLE rate decline hybrid model parameters are summarised in **Table 3**.

According to Temür et al.[35] the MSE and MAPE are used for ANN testing performance when considering simulated and real series. Hence, based on these performance parameters, the network model having 4 neurons/nodes in the hidden layer (1-4-1) was found to be the network with good performance (refer to Table 1 and the results highlighted in bold). The R^2 values for this network were 0.9743 and 0.9965 respectively, for the two sets of data.

Table 2: Summary for the ANN performance output for various numbers of neurons/nodes in the hidden layer for Canon Shale Well and Marcellus Shale Wells using JMP software

Canon Well				Marcellus Shale		
Number of Neurons	**MAPE**	**MSE**	**R^2**	**MAPE**	**MSE**	**R^2**
1	13.52	65.47	0.9727	13.18	182.07	0.9908
2	13.03	67.09	0.9726	42.92	116.11	0.9954
3	11.77	66.11	0.9730	54.11	141.94	0.9916
4	**7.96**	**10.49**	**0.9743**	**3.34**	**8.43**	**0.9965**
5	13.76	48.67	0.9855	26.95	86.96	0.9964
6	20.21	81.56	0.9667	27.94	7.08	0.9999
7	8.74	89.70	0.9864	13.28	15.00	0.9999
8	9.68	34.22	0.9886	4.82	4.68	1.0000
9	13.00	26.45	0.9930	7.89	6.91	1.0000
10	10.31	12.48	0.9961	3.30	3.06	1.0000
11	11.02	21.29	0.9958	21.25	31.36	0,9995
12	6.72	7.07	0.9982	9.78	6.17	1.000
13	6.82	9.26	0.9972	15.15	20.29	1.000
14	6.75	8.01	0.9978	14.73	26.02	1.000
15	11.06	28.48	0.9912	3.58	4.25	1.000

The ARIMA model of (3,1,3) showed good performance for the Canon Shale Well but failed to predict accurately for the Marcellus Shale Well i.e., MSE values of 1.2 and 3.6 respectively. The reason for the difference is that the model needs a large amount of production data to decide the best fitting model for a data series. Therefore the ARIMA model of (2,1,2) is recommended since it gave a good MSE and R^2 value for both wells (refer to Table 2 and the results highlighted in bold).

Table 3: The ARIMA model best fit and order determination for the Canon Shale Well and Marcellus Shale Well using JMP software

ARIMA	MSE (Canon Shale Well)	R^2 (Canon Shale Well)	MSE (Marcellus Shale Well)	R^2 (Marcellus Shale Well)
(0,0,0)	19.8	0.0000	66.7	0.0000
(1,1,1)	1.8	0.9880	2.4	0.9960
(1,2,1)	2.5	0.9780	2.7	0.9950
(1,3,1)	2.9	0.9670	3.8	0.9890
(2,1,1)	1.8	0.9880	2.5	0.9960
(2,1,2)	**1.6**	**0.9910**	**2.3**	**0.9960**
(3,1,3)	1.2	0.9940	3.6	0.9950

Table 3, summarises the behaviour, assumptions, condition and parameters for the single and hybrid decline models. The parameters were obtained using KAPPA-Citrine software.

Table 4: Summary of the behaviour, assumptions, condition and parameters for the decline models

	Canon Shale Well				Marcellus Shale Well			
Model	**Production Behaviour**	**Assumptions**	**Condition**	**Parameters**	**Production Behaviour**	**Assumptions**	**Condition**	**Parameters**
Arps	Linear to BDF	Decline parameter, b, defines the decline behaviour. The rate of decline approximated using a logistic growth decline	$0 \leq b \leq 1$ D is changing	$b = 0.383$ $D_i = 3.53$ $q_i = 2183$	BDF	Decline parameter, b, defines the decline behaviour. The rate of decline approximated using a logistic growth decline	$0 \leq b \leq 1$ D is changing	$b = 0.00$ $D_i = 0.042$ $q_i = 3864$
PLE	BDF and TFR	Estimates the rate of decline with a power law decline	D_i changes at early stages while D_∞ is constant at late time	$n = 0.524$ $D_i = 0.0000681$ $q_i = 4018$	BDF and TFR	Estimates the rate of decline with a power law decline	D_i changes at early stages while D_∞ is constant at late time	$n = 0.005$ $D_i = 0.0000431$ $q_i = 10812$
Arps-PLE hybrid decline model	BDF and TFR	Estimates the rate of decline with a logistic growth decline	$0 \leq b \leq 1$ D_i changes at initial stages and then becomes constant at late time	$n = 1.00$ $D_i = 0.798$ $q_i = 1429$ $b = 0.000$	BDF and TFR	Estimates the rate of decline with a logistic growth decline	$0 \leq b \leq 1$ D_i changes at initial stages and then becomes constant at late time	$n = 0.005$ $D_i = 0.0139$ $q_i = 359$ $b = 0.0087$

Figure 3 and 4 displays the scatter graph of the forecasted shale gas production data by the Arps decline model compared with the actual data for the Canon Shale Well and Marcellus Shale Well, respectively.

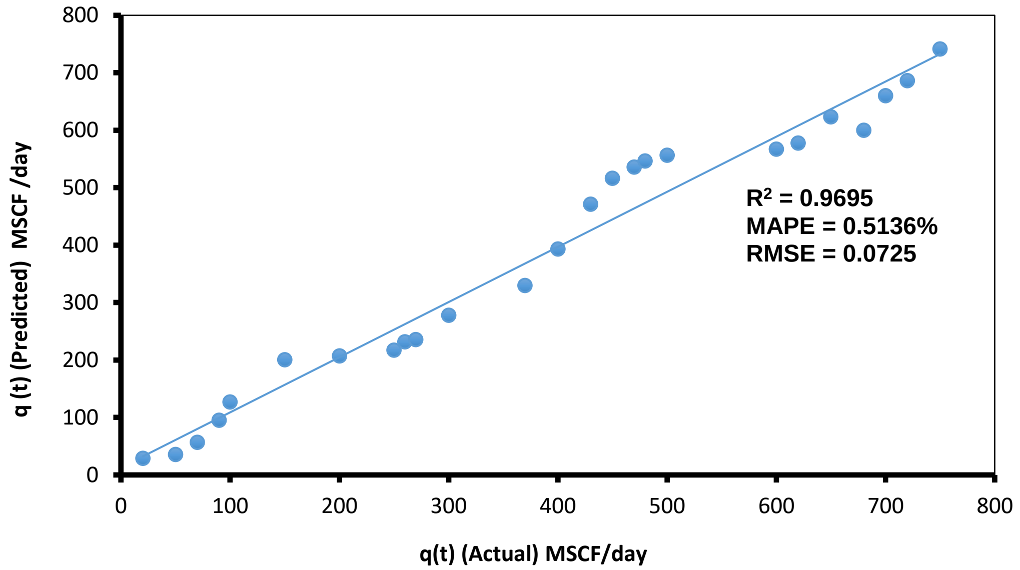


Figure 3: Scatter graph of the forecasted shale production data by the Arps decline model vs. actual data for Canon Shale Well

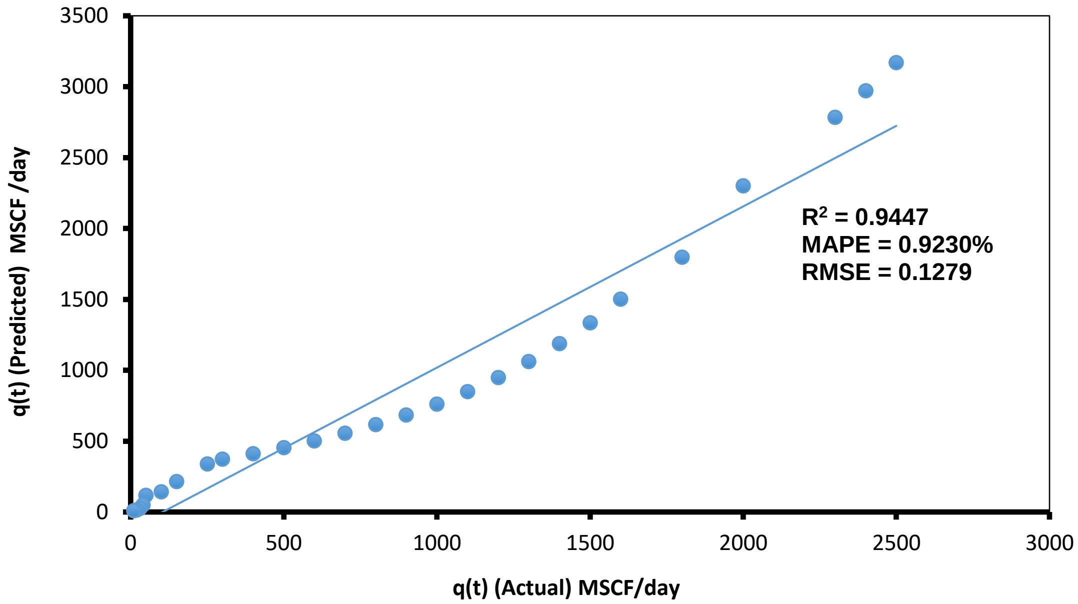


Figure 4: Scatter graph of the forecasted shale production data by the Arps decline model vs. actual data for Marcellus Shale Well

Figures 3 and 4 illustrations the R^2 figures of 0.9695 and 0.9447 obtained for the Canon and Marcellus Shale Wells respectively. While a MAPE of 0.5136 and 0.9230%

were obtained for the two sets of production data. The RMSE figures were 0.0725 and 0.1279 respectively. The results for the model appear in some instances to over-and in other instances to under-estimate the data. The results concur with the literature, which suggests that the weakness of the Arps model is over-estimation of results. Tan et al.[15] in their study highlighted that while the Arps decline is simplistic and fast, it regularly fails to precisely fit the decline curve of SGRs. They further explained that the model frequently inclines to overestimate the EUR as it assumes that a BDF is evident. Paryani et al.[8] concurred with these findings, explaining that the drainage area is not constant because the pressure pulse continues to spread from the fracture to other areas of the reservoir volume. Under these conditions, the b value predicted by the Arps model for the actual production data will be larger than 1. This in turn leads to inaccurate estimations of reserves i.e. the underestimation of results.

Figure 5 and 6 displays the scatter plot of the predicted shale gas production data by the PLE decline model compared to the actual data for the Canon Shale Well and Marcellus Shale Well, respectively.

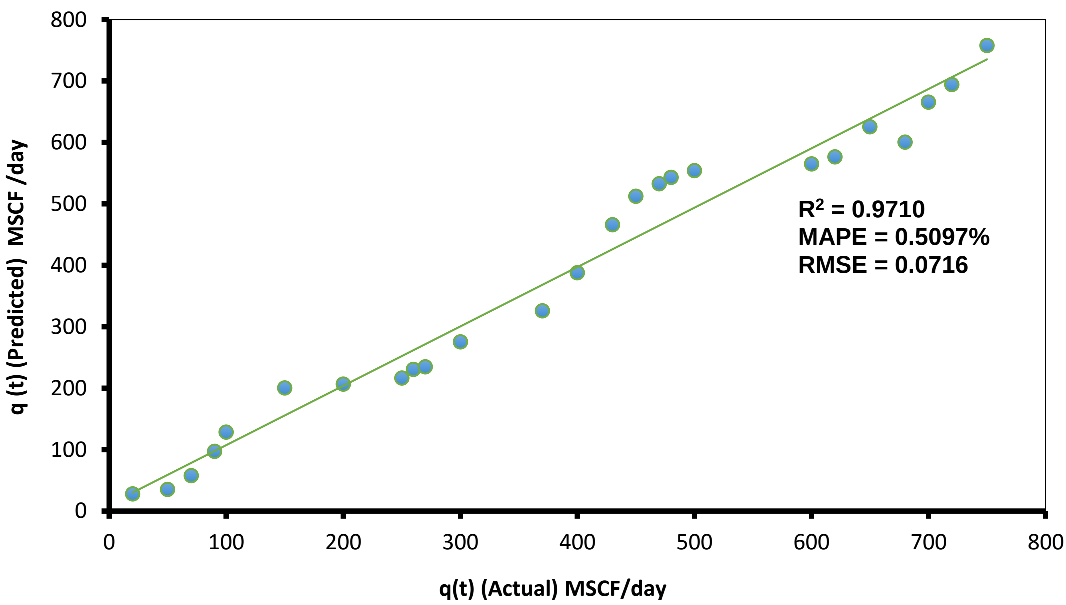


Figure 5: Scatter graph of the forecasted shale production data by the PLE decline model vs. actual data for Canon Shale Well

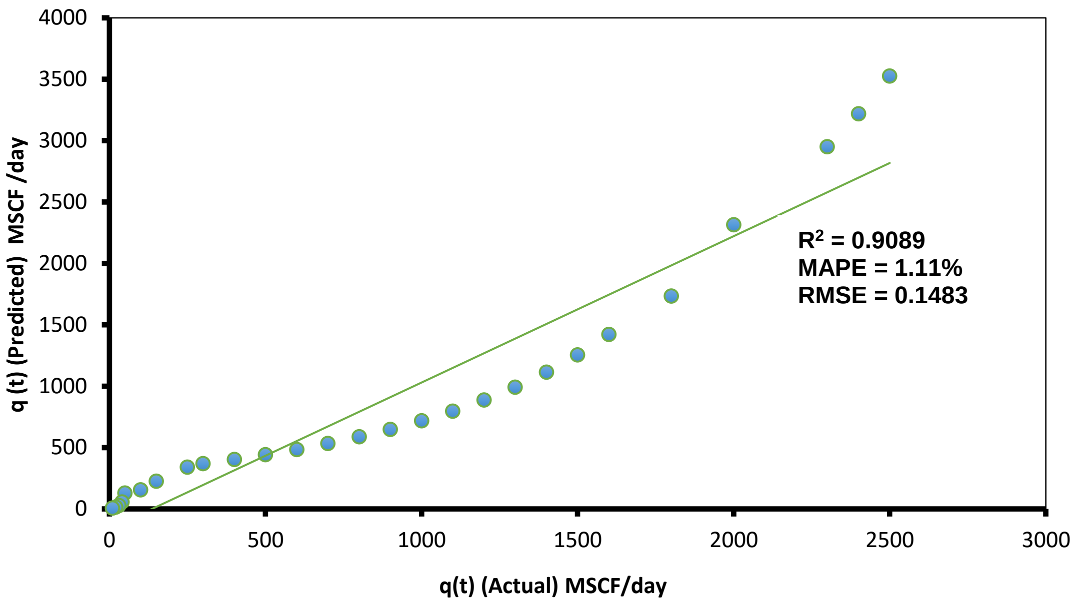


Figure 6: Scatter graph of the forecasted shale production data by the PLE decline model vs. actual data for Marcellus Shale Well

Figures 5 and 6 shows the R^2 values of 0.9710 and 0.9089 were obtained for the Canon and Marcellus Shale Wells, respectively. While a MAPE of 0.6097 and 0.1.11% was obtained for the two sets of production data. The RMSE figures were 0.0716 and 0.1483, respectively. The results for the PLE model appear to underestimate the data although the PLE considers BDF and TFR, which is an advantage of the model. Furthermore, the model was specifically developed for SGRs; hence it was assumed that the results would be better. The underestimation in the results concurred with the findings by Paryani et al.[8] as, based on their results, the PLE frequently offered the lowest forecasts. The model is thus the most conservative method for production forecasting and reserves estimation.[8] Seshadri et al.[36] determined that for tight gas wells, the PLE model is complex and non-intuitive. The PLE model can result in a non-unique solution due to four degrees of freedom, resulting from the four unknown parameters, this may contribute to the underestimation of the results.

Figure 7 and 8 displays the scatter plot of the predicted shale gas production data by the established ARIMA model compared to the actual data for the Canon Shale Well and Marcellus Shale Well, respectively.

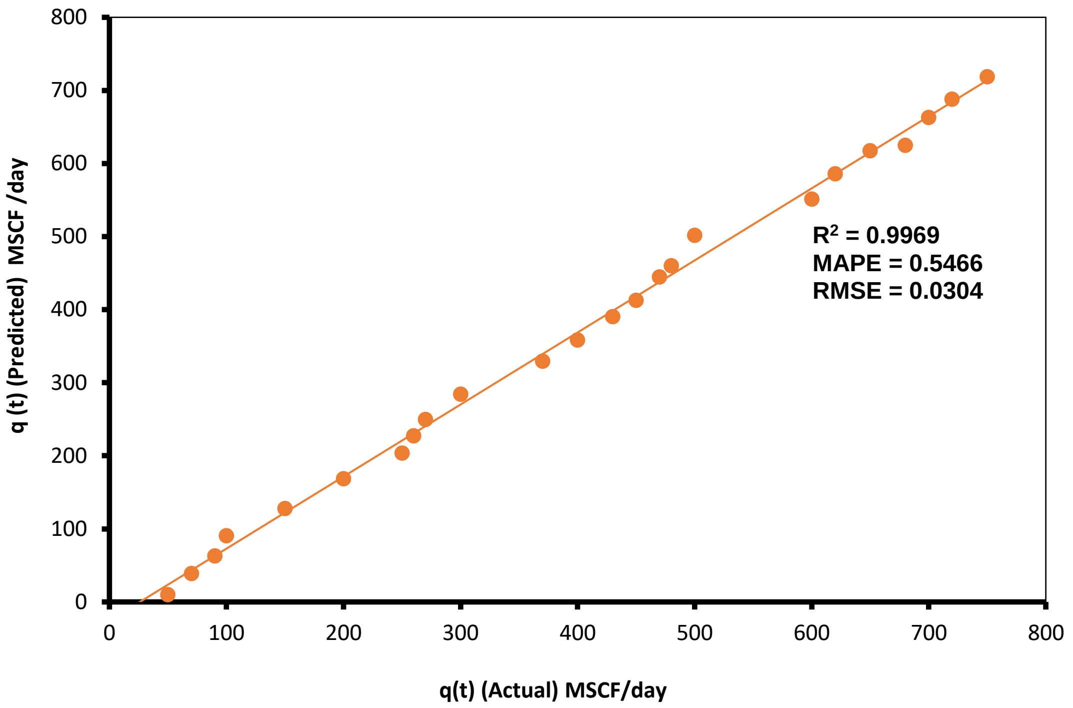


Figure 7: Scatter graph of the forecasted shale production data by the ARIMA model vs. actual data for Canon Shale Well

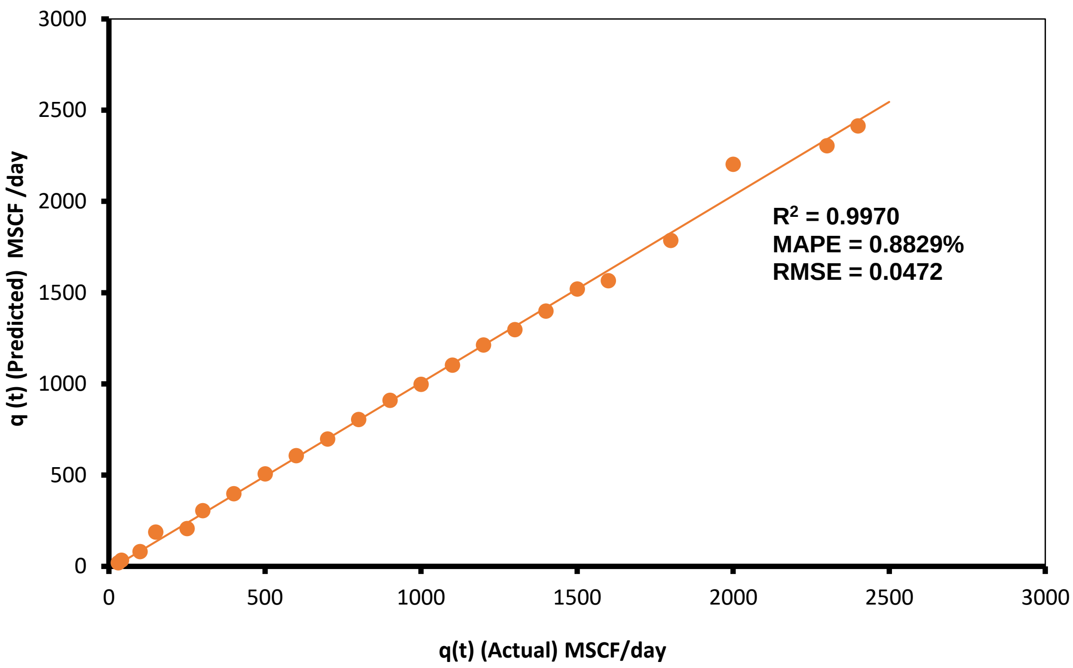


Figure 8: Scatter graph of the forecasted shale production data by the ARIMA model vs. actual data for Marcellus Shale Well

From **Figure 7 and 8**, R^2 values of 0.9969 and 0.9970 were obtained for the Canon and Marcellus Shale Wells, respectively. While a MAPE of 0.5466 and 0.8829% were obtained for the two sets of production data. The RMSE numbers were 0.0304 and 0.0472, respectively. The expected outcomes from the model appear to follow a close trend to the actual values according to the $R^{2.}$ However, the model gave higher MAPE values when compared to the ANN model. Reikard et al.[37] suggested that ARIMA models have shown to be outstanding interim forecasting models for a wide variety of time series because interim factors are anticipated to change slowly.

Figure 9 and 10 displays the scatter plot of the predicted shale gas production data by the established ANN model when compared to the actual data for the Canon Shale Well and Marcellus Shale Well, respectively.

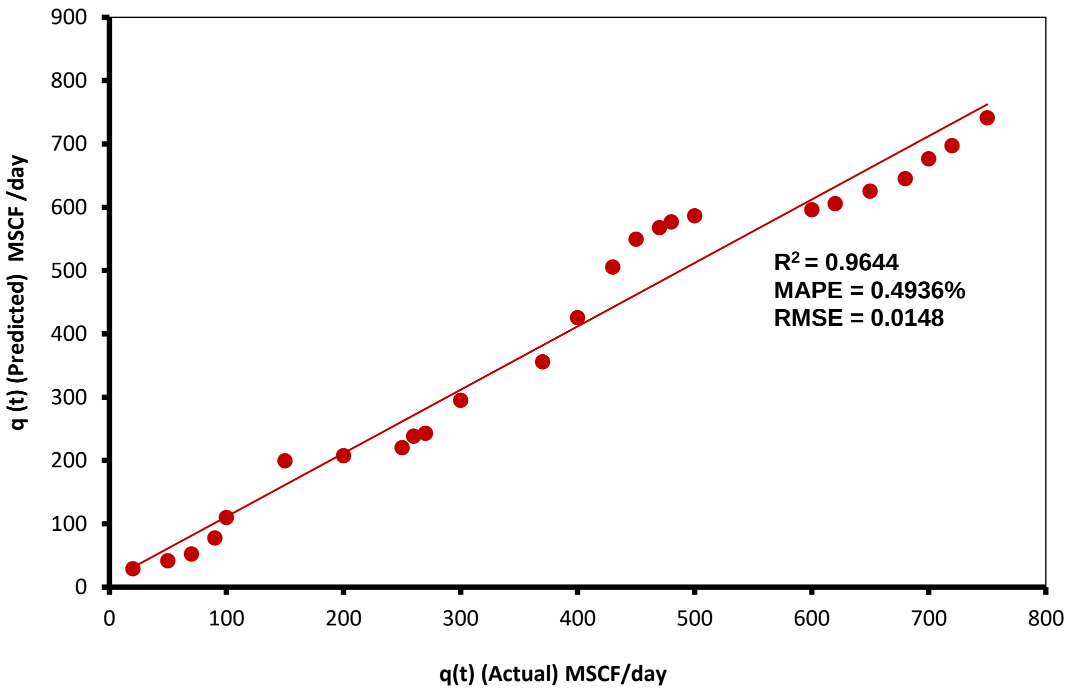


Figure 9: Scatter graph of the forecasted shale production data by the ANN model vs. actual data for Canon Shale Well

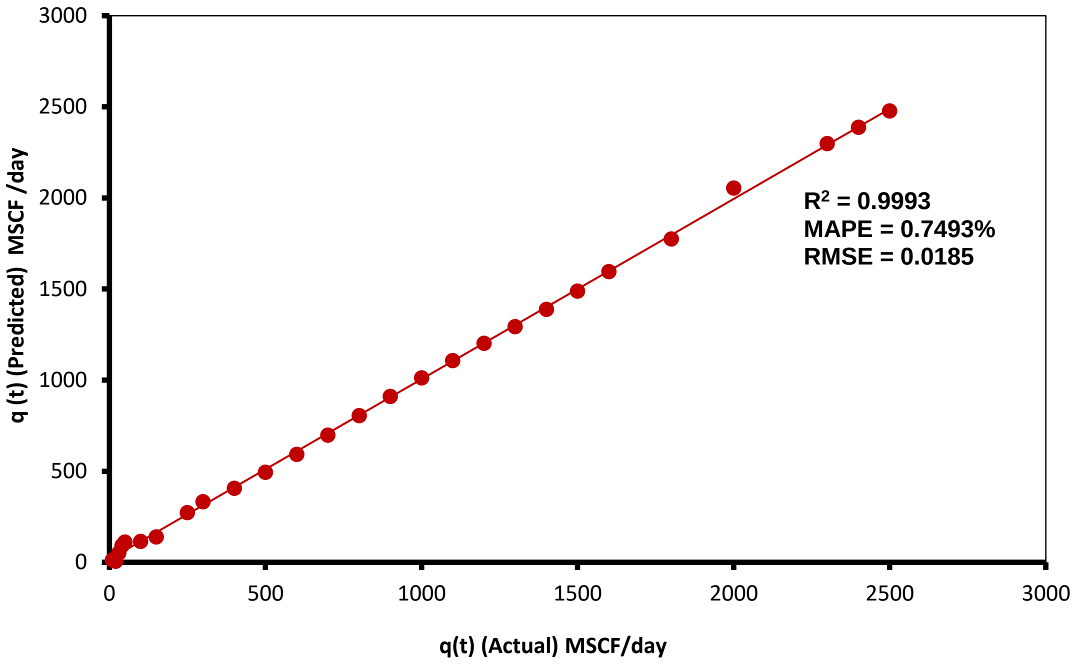


Figure 10: Scatter graph of the forecasted shale production data by the ANN model vs. actual data for Marcellus Shale Well

As can be seen from **Figure 9 and 10**, R^2 values of 0.9644 and 0.9993 were obtained for the Canon and Marcellus Shale Wells respectively. While a MAPE of 0.4946 and 0.7493% was obtained for the two sets of production data, the RMSE values were 0.0148 and 0.0185, respectively. Lai et al.[38] indicated that neural networks are useful for modelling and predicting the properties of time series data. Dhini et al.[32] described neural networks as having a universal non-linear function and a relatively high degree of forecasting accuracy. In addition, neural network forecasting provides better results than traditional forecasting methods, over monthly as well as quarterly periods.[32]

Figure 11 and 12 displays the scatter plot of the predicted shale gas production data by the ANN-ARIMA hybrid model compared to the actual data for the Canon Shale Well and Marcellus Shale Well, respectively.

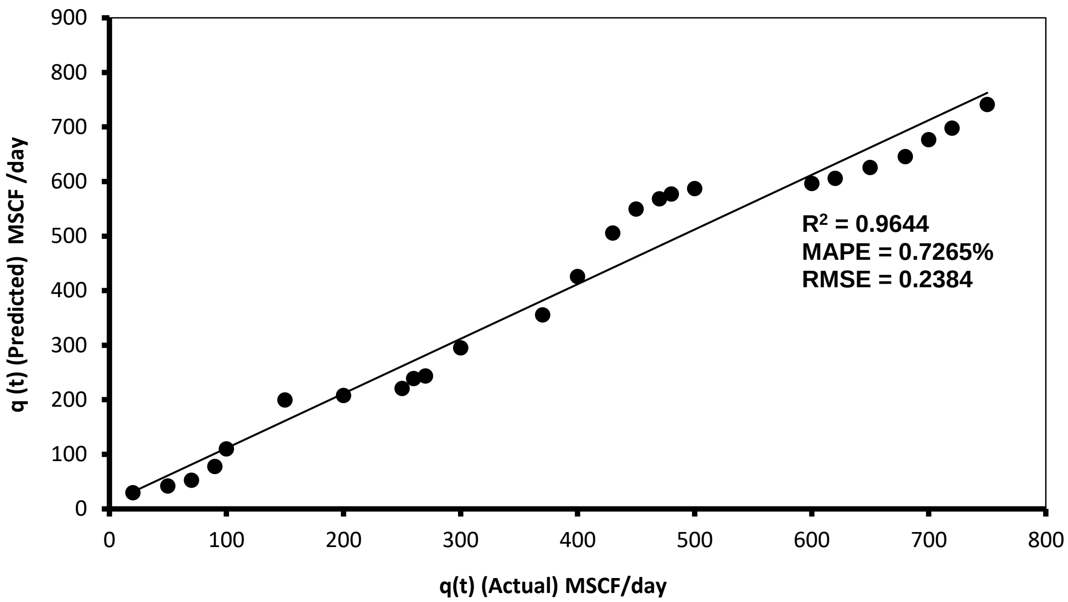


Figure 11: Scatter graph of the forecasted shale production data by the ANN-ARIMA hybrid model vs. actual data for Canon Shale Well

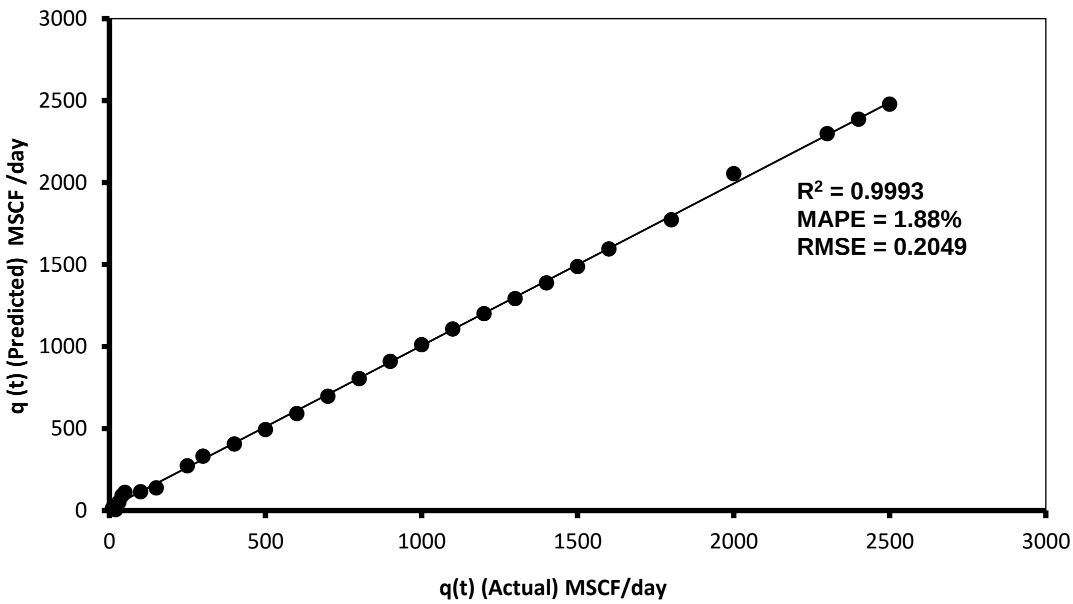


Figure 12: Scatter graph of the forecasted shale production data by the ANN-ARIMA hybrid vs. actual data for Marcellus Shale Well

Figures 11 and 12 shows the R^2 values of 0.9644 and 0.9993 were obtained for the Canon and Marcellus Shale Wells, respectively. While a MAPE of 0.7265 and 1.88% was obtained for the two sets of production data. The RMSE values were 0.2384 and 0.2049 respectively. This result appears to have a higher error than ANN and ARIMA, which is contradictory to the findings in the literature. According to Faruk[30], hybrid approaches have a higher degree of accuracy than neural networks. Dhini et al.[32] indicated that hybrid models bring together the benefits of ARIMA and ANN. However, Taskaya-Temizel et al.[6] stated that under certain conditions, the single model approach can outdo hybrid models. This was observed during this study.

Figure 13 and 14 displays the scatter plot of the predicted shale gas production data by the established Arps-PLE hybrid decline model compared to the actual data for the Canon Shale Well and Marcellus Shale Well, respectively.

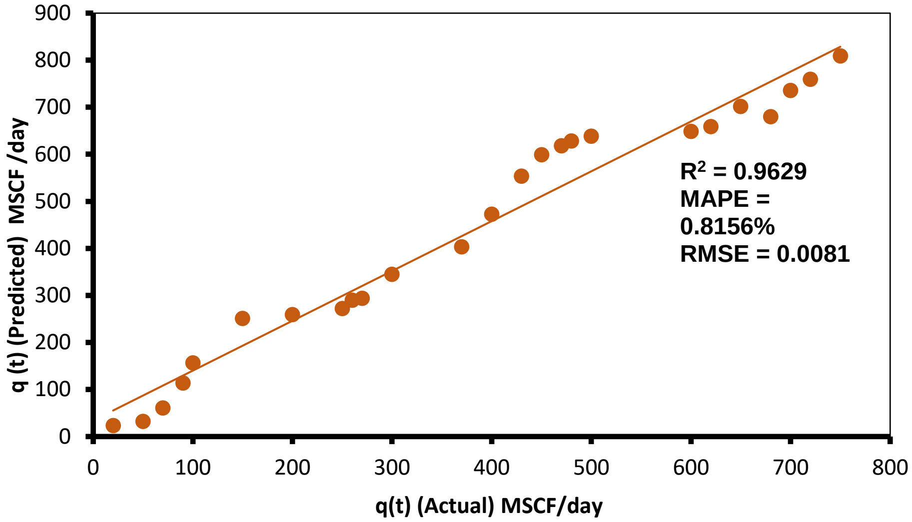


Figure 13: Scatter plot of the predicted shale production data by the Arps-PLE hybrid decline model vs. actual data for Canon Shale Well

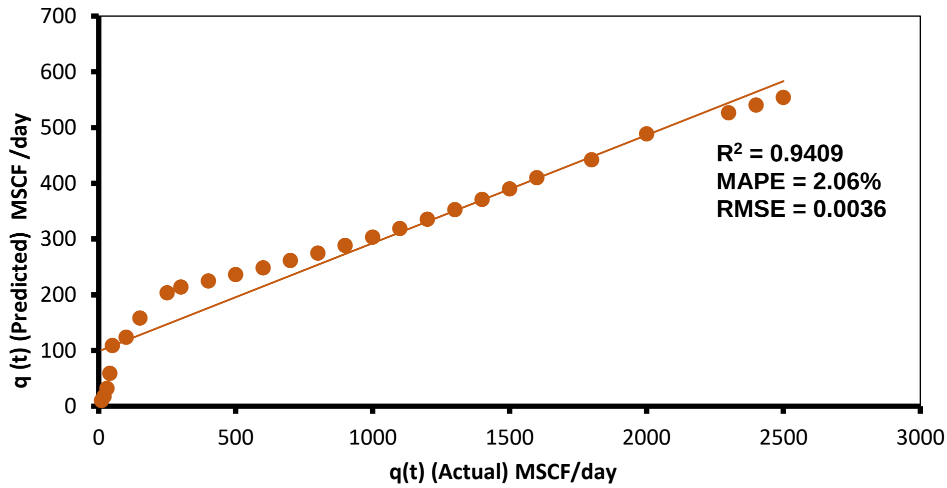


Figure 14: Scatter plot of the predicted shale production data by the Arps-PLE hybrid decline model vs. actual data for Marcellus Shale Well

From **Figure 13 and 14**, R^2 values of 0.9629 and 0.9409 were obtained for the Canon and Marcellus Shale Wells, respectively. While a MAPE of 0.8156 and 2.06% were obtained for the two sets of production data. The RMSE figures were 0.0081 and 0.0036, respectively. The results from the model initially appear to over-and under-estimate the data prediction. However, the results tended to move closer to the actual values over time. This would be attributed to the reliability of the Arps model and the fact that the PLE model was developed precisely for SGR. Moreover, both flow regimes are considered and subsequently most shale gas wells seldom reach BDF. The results appear to move closer to the actuals when approaching the TFR. Hence, by combining the models, the overestimation of the predicted results is minimised over time.

Table 5, summaries the RSME, MAPE and R^2 for the models which are ranked in order of most to least accurate, based on the RSME number. The lower the RMSE value is, the better the model. Also, the RMSE is more accurate than the MAPE value.

Table 5: **Summary of the accuracy and validation in ranking order, from most to least for the different models**

	Canon Shale Well			**Marcellus Shale Well**		
	RMSE	**MAPE (%)**	**R^2**	**RMSE**	**MAPE (%)**	**R^2**
Arps-PLE Hybrid Model	0.0081	0.8156	0,9629	0.0036	2.0634	0,9409
ANN Model	0.0148	0.4936	0,9644	0.0185	0.7493	0,9993
ARIMA Model	0.0304	0.5466	0.9969	0.0472	0.8829	0.9970
PLE Decline Model	0.0716	0.5097	0,9710	0.1483	1.1102	0,9089
Arps Decline Model	0.0725	0.5136	0.9695	0.1279	0.923	0,9447
ANN-ARIMA Hybrid Model	0.2304	0.7265	0.9644	0.2049	1.8837	0.9993

Based on the summary results presented in **Table 5**, it is evident the Arps-PLE hybrid decline model is a better model than the other models, followed by the ANN and

ARIMA models, based on the RMSE value. The ANN-ARIMA model appears to accurately predict the results compared to the Arps-PLE hybrid decline model i.e. Figure 12 vs. Figure 14, if the R^2 value is considered. However, one needs to keep in mind, that the R^2 denotes relative degree of fit, while RMSE indicates absolute degree of fit and should be used as a determining factor when considering a model i.e. the RMSE value.

It is apparent that by combining the Arps and PLE decline models, the limitations of the models is reduced. There is a significant reduction in the RMSE values from 0.0716 and 0.0725 for the PLE and Arps models, respectively; to 0.0081 for the Arps-PLE hybrid decline model, for the Canon Shale Well; while for the Marcellus Shale Well, 0.1483 and 0.1279 to 0.0036. The contributing factor that was highlighted earlier is the dominance of the PLE parameters in the model i.e. D_i and D_∞ which considers the TFR flows not catered for in the Arps decline model. In addition, the PLE model was specifically developed for SGRs and by combining the models, the number of variables are reduced from four to three. This reduces the number of variables to solve which was identified as a limitation of the model and hence the degrees of freedom is reduced. After the validation of the models, a step was taken further, and the 95% confidence intervals were assessed. **Figures 15 and 16**, illustrates the findings for analysis.

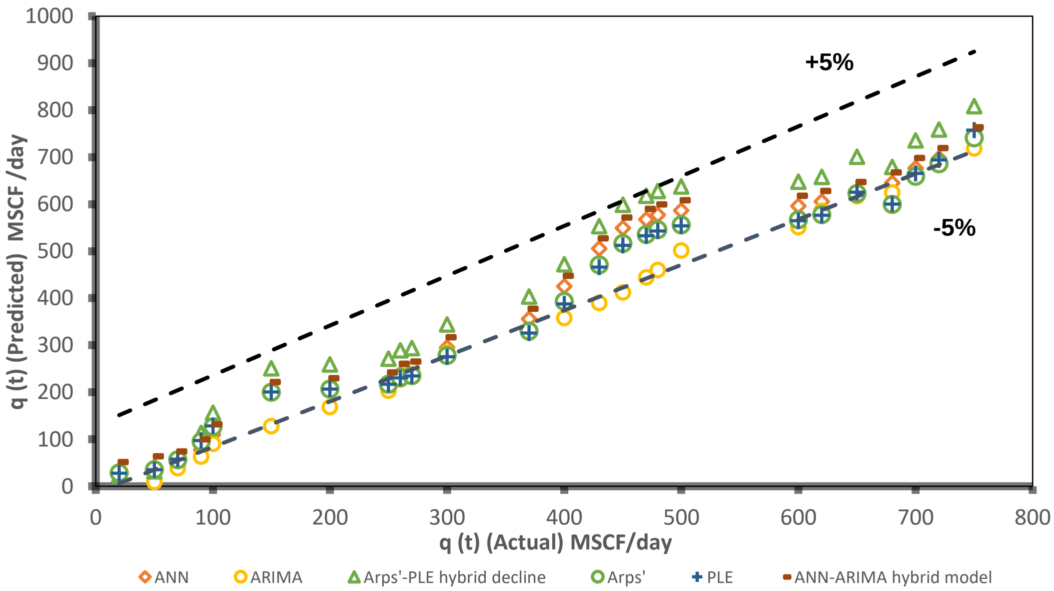


Figure 15: Scatter graph of the forecasted shale production data by the ANN and ARIMA models vs. actual data for Canon Shale Well; evaluating the 95% confidence level

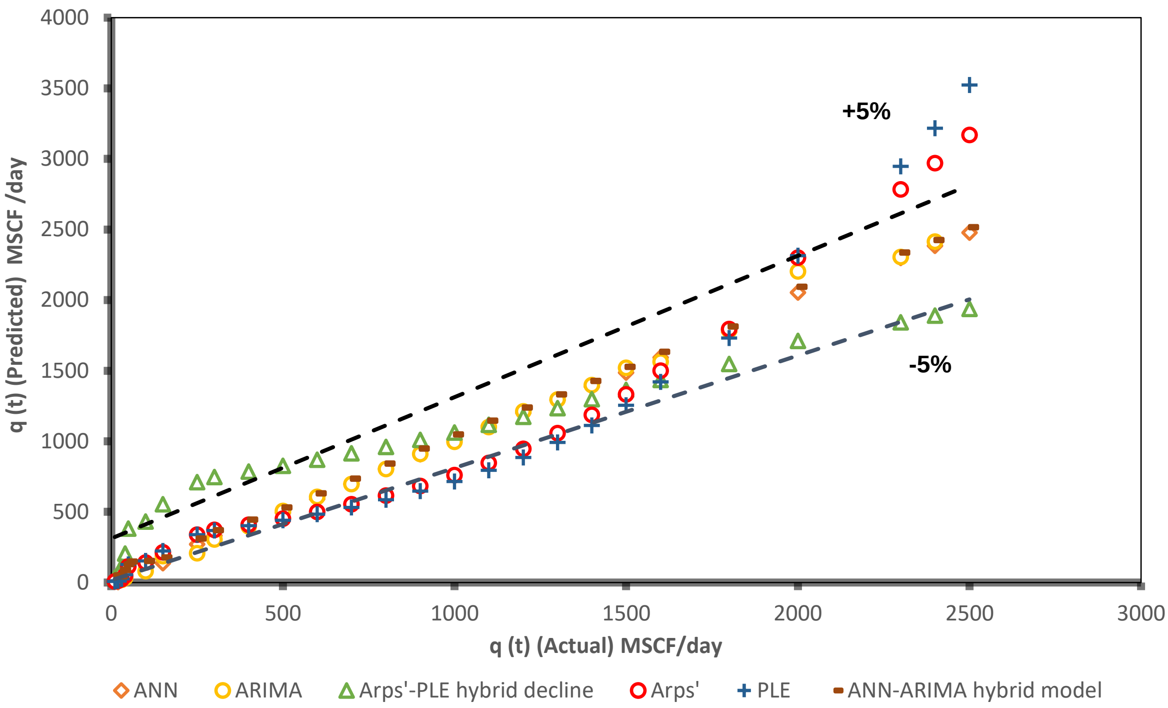


Figure 16: Scatter graph of the forecasted shale production data by the ANN and ARIMA models vs. actual data for Marcellus Shale Well evaluating the 95% confidence level

Evaluating the confidence intervals is one way to aid to evaluate what the values potentially might be in the wider population.[39] Confidence intervals offer a possible array of values, bracketed by lower and upper limits that encompass the unknown population or 'true' value expected by that sample mean, correlation coefficient or odds ratio.[39] It is standard to report either the 90%, 95% or 99% confidence interval (CI); the 95% CI tends to be commonly used.[38] From **Figures 15 and 16**, it is clearly evident that for the ANN, ARIMA, ANN-ARIMA hybrid and Arp's-PLE hybrid decline models are confidently within the 95% limit i.e. the data lies within the ±5% error band for the two data sets. However, the other models showed inconsistency between the sets of data i.e. in one instance within the range and in another outside. The results show very good fit between the actual and forecast values for the ANN, ARIMA, ANN-ARIMA hybrid and Arp's-PLE hybrid decline models.

5. Conclusions

A multilayer perception (MLP) ANN and ARIMA model was established to forecast shale gas production. The best fitting data was obtained with a 1-4-1 structure for the

ANN model, while with the ARIMA model, the number of nodes which showed the best performance indicated a (3,1,3) and (2,1,2) for the Canon Well Shale and Marcellus Shale, respectively. A MAPE of 7.96 and 3.34% was identified to be the network with the best performance while the R^2 values for this network were 0.9743 and 0.9965, respectively for the two sets of data. The ARIMA performance evaluation of (3,1,3) and (2,1,2) was found to have the lowest MSE of 1.20 and 2.30, while the adjusted R^2 was 0.9940 and 0.9960 for the Canon Shale Well and Marcellus Shale Well respectively. Also, a hybrid decline model was developed by combining the Arps and PLE decline models. The developed hybrid model was then evaluated and compared with the Arps, PLE, ANN, ARIMA and ANN-ARIMA hybrid models. The following can be concluded:

- The current DCA models, Arps and PLE decline models appear to over and under-estimate the data.
- The Arps-PLE hybrid decline model gave the closest predicted production data, compared to the ANN-ARIMA hybrid model.
- Both the ARIMA and ANN models gave good predicted results. However, when both models were combined into the ANN-ARIMA hybrid model, the strengths of both models referenced in literature did not provide accurate predictive data. The result was an overestimation in the production flow rate.
- The validation of the models indicated that the Arps-PLE hybrid decline model gave the lowest RMSE value with a good R^2 value for both the Canon Shale Well and Marcellus Shale Well; followed by the ANN and ARIMA models.
- Lastly, the confidence interval evaluation found that the ANN, ARIMA, ANN-ARIMA hybrid and Arp's-PLE hybrid decline models fell within the 95% confidence limit, i.e. the data lies within the ±5% error band. The results proved that there was a very good fit between the actual and forecasted values for the models.

In conclusion, the findings have provided a significant contribution to the forecasting of shale gas production. The results indicated that the Arps-PLE hybrid decline model is a good model predictor for shale gas production. The contributing factor is the dominance of the PLE parameters i.e. D_i changes in the early stages and D_∞ becomes constant at later times in the model. This caters for the TFR which the Arps decline model did not consider. Lastly, with the PLE the limitation identified during the sensitivity analysis was the number of variables, so, by combining the models the variables is reduced and hence the degrees of freedom is reduced from four to three.

Future research will investigate incorporating and accounting for the Knudsen diffusion into the Arps-PLE hybrid decline model in shale gas reservoir modelling.

References

1. Zhang, X.; Wang, X.; Hou, X.; Xu, W. Rate decline analysis of vertically fractured wells in shale gas reservoirs. *Energies*. **2017**, *10*(10), 1602.
2. Zhang, H. E.; Nelson, E.; Olds, D.; Rietz, D.; Lee, W. J. Effective Applications of Extended Exponential Decline Curve Analysis to both Conventional and Unconventional Reservoirs. In SPE Annual Technical Conference and Exhibition, Dubai, UAE, 26-28 September 2015.
3. Manda, P.; Nkazi, D. B. The Evaluation and Sensitivity of Decline Curve Modelling. *Energies*. **2020a**, *13*(11), 2765.
4. Alabboodi, M. J. Conditioning the Estimating Ultimate Recovery of Shale Wells to Reservoir and Completion Parameters. In Proceedings of the SPE Eastern Regional Meeting, Canton, OH, USA, 13-15 September 2016.
5. Manda, P.; Nkazi, D. B. Production Decline Prediction of Shale Gas using Hybrid Models. *GJRE*. **2020**, *20*(5), 1-15.
6. Taskaya-Temizel, T.; Ahmad, K. Are ARIMA neural network hybrids better than single models? Proceeding of International Joint Conference on NNs, Montreal, Canada, 31 July-4 August 2015.
7. Boah, E. A.; Borsah, A. A.; Brantson, E. T. Decline Curve Analysis and Production Forecast Studies for Oil Well Performance Prediction: A Case Study of Reservoir X. *Int. J. Eng. Sci.* **2018**, *7*, 56-67.
8. Paryani, M.; Ahmadi, M.; Awoleke, O.; Hanks, C. Decline curve analysis: A Comparative Study of Proposed Models Using Improved Residual Functions. *J. Pet. Environ. Biotechnol*. **2018**, *9*(1), 1-8.
9. Ali, T. A.; Sheng, J. J. Production decline models: A comparison study. In SPE Eastern Regional Meeting, Morgantown, WV, USA, 13-15 October 2015.
10. Yuhu, B.; Guihua, C.; Bingxiang, X.; Ruyong, F.; Ling, C. Comparison of typical curve models for shale gas production decline prediction. *CN. Pet. Explor*. **2016**, *21*(5), 96-102.
11. Li, P.; Hao, M.; Hu, J.; Ru, Z.; Li, Z. (2018). A new production decline model for horizontal wells in low-permeability reservoirs. *J. Pet. Sci. Eng.* **2018**, *171*, 340-352.
12. Arps, J. J. Analysis of decline curves. *Transactions of the AIME*. **1945**, *160*(1), 228-247.

13. Qu, Z. *Proceedings of the International Field Exploration and Development Conference 2017*; Lin, J.E., Ed.; Springer: Singapore, 2018; pp 75-76.
14. Ilk, D.; Rushing, J. A.; Blasingame, T. A. Integration of production analysis and rate-time analysis via parametric correlations--theoretical considerations and practical applications. In SPE Hydraulic Fracturing Technology Conference and Exhibition, Woodlands, Texas, 24–26 January 2011.
15. Tan, L.; Zuo, L.; Wang, B. Methods of decline curve analysis for shale gas reservoirs. *Energies*. **2018**, *11*(3), 552.
16. Kanfar, M.; Wattenbarger, R. Comparison of empirical decline curve methods for shale wells. In Proceedings of the SPE Canadian Unconventional Resources Conferences, Calgary, AB, Canada, 30 October-1 November 2012.
17. Vanorsdale, C. R. Production decline analysis lessons from classic shale gas wells. In SPE Annual Technical Conference and Exhibition, New Orleans, LA, USA, 30 September- 2 October 2013.
18. Hu, Y.; Weijermars, R.; Zuo, L.; Yu, W. Benchmarking EUR estimates for hydraulically fractured wells with and without fracture hits using various DCA methods. *J. Pet. Sci. Eng*. **2018**, *162*, 617-632.
19. Johnson, N. L.; Currie, S. M.; Ilk, D.; Blasingame, T. A. A simple methodology for direct estimation of gas-in-place and reserves using rate-time data. In SPE Rocky Mountain Petroleum Technology Conference, Denver, CO, April 2009.
20. Jackson, E. A.; Sillah, A.; Tamuke, E. Modelling monthly headline consumer price index (HCPI) through seasonal Box-Jenkins methodology. *Int. J. Sci*. **2018**, *7*(1), 51-56.
21. Fathi, O. Time series forecasting using a hybrid ARIMA and LSTM model. In Velvet Consulting, France, 2019.
22. Fattah, J.; Ezzine, L.; Aman, Z.; El Moussami, H.; Lachhab, A. Forecasting of demand using ARIMA model. *IJEBM*. **2018**, *10*, 1-9.
23. Lari, S. Assessment of Geometrical Features of Internal Flaws with Artificial Neural Network Optimized by a Thermodynamic Equilibrium Algorithm. M.Sc. Thesis, University of Waterloo, 2020.
24. Dimla, E. Artificial Neural Networks: A Manufacturing Engineering Perspective and Case Study Review. *J. Commun*. **2019**, *14*(8).
25. Du, K. L.; Swamy, M. N. Neural networks in a soft computing framework. Springer-Verlag: London, 2006, 3-566.

26. Azizi, S.; Ahmadloo, E. Prediction of heat transfer coefficient during condensation of R134a in inclined tubes using artificial neural network. *Appl. Therm. Eng*. **2016**, *106*, 203-210.
27. Ross, M.; Berberian, N.; Chartier, S. Should I Stay or Should I Grow? A Dynamic Self-Governed Growth for Determining Hidden Layer Size in a Multilayer Perceptron. In International Joint Conference on Neural Networks, Canada, 19 July 2020.
28. Kuila, U.; Prasad, M.; Kazemi, H. Assessing Knudsen flow in gas-flow models of shale reservoirs. *Can. Soc. Explor. Geophys*. **2013**, *38*, 23-27.
29. Huang, T.; Guo, X.; Wang, K. Nonlinear seepage model of gas transport in multiscale shale gas reservoirs and productivity analysis of fractured well. *J. Chem*. *2015*, *1*, 349-359.
30. Faruk, D. Ö. A hybrid neural network and ARIMA model for water quality time series prediction. *Eng. Appl. Artif. Intell*. **2010**, *23*(4), 586-594.
31. Zhang, M.; Li, J.; Pan, B.; Zhang, G. Weekly Hotel Occupancy Forecasting of a Tourism Destination. *Sustainability*. **2018**, *10*(12), 4351.
32. Dhini, A.; Surjandari, I.; Riefqi, M.; Puspasari, M. A. Forecasting analysis of consumer goods demand using neural networks and ARIMA. *Int. J. Technol*. **2015**, *6*(5), 872-880.
33. Chao, M. A.; Kulkarni, C.; Goebel, K.; Fink, O. Hybrid deep fault detection and isolation: Combining deep neural networks and system performance models. *Int. J. Progn. Health Manag*. **2019**, *10* (11), 1-25.
34. Granger, C. W. J. Combining Forecasts—Twenty Years Later. *Essays in Econometrics: Collected Papers of Clive WJ Granger*, **2001**, *32*, 411.
35. Temür, A. S.; Akgün, M.; Temür, G. Predicting housing sales in Turkey using ARIMA, LSTM and hybrid models. *J. Bus. Econ. Manag*. **2019**, *20*(5), 920-938.
36. Seshadri, J. N.; Mattar, L. Comparison of power law and modified hyperbolic decline methods. In Canadian Unconventional Resources and International Petroleum Conference, Calgary, AB, Canada, 19-21 October 2010.
37. Reikard, G.; Robertson, B.; Bidlot, J. R. Combining Wave Energy with Wind and Solar: Short-term Forecasting. *Renew. Energ*. **2015**, *81*, 442-456.
38. Lai, G., Chang, W. C., Yang, Y., & Liu, H. Modelling long-and short-term temporal patterns with deep neural networks. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. **2018**, 94-104.

39. Fethney, J. Statistical and Clinical Significance, and How to Use Confidence Intervals to Help Interpret Both. *Au. Crit. Care*. **2010**, *23*(2), 93-97.