**Title:** **Utilisation of Artificial Intelligence based Time-Series Prediction to validate Carbon Containment in Injection Well in Illinois Basin**

**Authors:** Munish Kumara and Kanna Swaminathana

**Affiliations:** aSingapore University of Social Sciences, School of Business, Clementi Rd, #463号, Singapore 599494

bCalifornia Resources Corporation, 900 Old River Rd, Bakersfield, CA 93311, United States

Corresponding author. Email: munishkumar001@suss.edu.sg

**Abstract:**

**One-Sentence Summary:** Applying machine learning and predicative analytics via time series injection information and monitoring data on a carbon capture well to predict well injection rate deltas.

**Keywords (minimum 6):** low-carbon, time-series, neural network, LSTM, carbon capture, injection pressure, monitoring, plume migration

# Introduction

There is no doubt that human activities are one of the main reasons for the increase in amounts of greenhouse gases (GHG) and global warming. Primarily, carbon dioxide (CO2) levels have increased due to the rapid pace of industrialisation and population growth, which has only picked up pace since the 1960s [1].

The burning of fossil fuels (coal, oil, and natural gas) has resulted in carbon dioxide (CO2), methane (CH4) and nitrous oxide (N2O) generation. A report generated by the United States Environmental Protection Agency (EPA) evaluated the amount of CO2 generated in 2020 to be ~3.11 million metric tons [2]. Such significant quantities, if not better managed, will lead to irreversible environmental consequences.

Carbon Capture Utilisation and Storage (CCUS) is one promising method to deal with the produced CO2 from anthropogenic sources. The idea is to capture CO2 from an emission point source and subsequent sequester it via injection into a suitable geological formation, with the explicit aim to store the CO2 safely, in a state of permanence. CCUS is now being viewed as a key technology to assist in reaching global anthropogenic climate change goals. Typically, CO2 storage is carried out in one of three ways - via injection into virgin saline aquifers; into depleted oil and gas fields; or if used for Enhanced Oil Recovery (EOR) processes. These methods all have different project drivers, risks, and commercial implications. However, what these 3 methods have in common is the requirement that (a) there be a good understanding of the subsurface geological properties and (b) there be some ability to monitor and even predict CO2 behaviour at the well scale, be in during the injection phase or during the shut-in phase.

The uncertainties present in (b) especially necessitate pilot development and validation of the technology. The Illinois Basin - Decatur Project is one such study which aimed to demonstrate the capacity, injectivity, and containment of carbon storage in the Mount Simon Sandstone, which is the main carbon storage resource in the Illinois Basin and the Midwest Region. The source of the injected CO2 is from ethonal production at the Archer Daniels Midland company’s plant. The CO2 is compressed, dehydrated and injected into the Mt. Simon Sandstone, which is primarily a saline aquifer approximately ~7,000 ft deep. Injection began in 2009 and continued for a 3-year period (Nov-2011 to Nov-2014). Cumulatively, ~999,215 tonnes of supercritical CO2 have been injected and geologically stored. An injection and verification well, ~700ft apart, was drilled into the formation. The wells were equipped with downhole sensors to monitor pressure and temperature at various depths of interest.

|  |
| --- |
|  |

Figure 1: Location map of the Illinois Basin – Decatur Project (IBDP). Image taken from

This paper aims to use time series injection information and monitoring data on a carbon capture well to predict carbon capture well injection rates deltas (D) which is the difference in the injection rate (IR) at time t and time (t-1) i.e.

|  |  |  |
| --- | --- | --- |
|  |  | Equation 1 |

Correlating the change in injection rate to the behaviour of other parameters in the well can be used to provide a checkpoint against carbon migration from the well or other losses during the process. Utilisation of a machine learning (ML) method to predict injection rate deltas based on monitoring well data can be used to validate carbon containment throughout the injection of the well as well.

# Literature Review

Various authors have tried numerous methods to forecast future trends based on past data. Work by De Gooijer and Hyndman [3], for instance, reviewed a series of time-series forecast models over a 25 year period, from 1985 to 2005. Their review highlighted various models being developed and applied in a myriad of scenarios related to finance, statistics and manufacturing, and include methods such as (a) exponential smoothing [4, 5], (b) Autoregressive Integrated Moving Average (ARIMA) [6], (c) seasonal models [7], (d) state space and structural models and the Kalman filter [8], (e) nonlinear models [9], (f) long-range dependence models, including the family of Autoregressive Fractionally Integrated Moving Average (ARFIMA) models [10], (g) Autoregressive Conditional Heteroskedasticity/Generalized Autoregressive Conditional Heteroskedasticity (ARCH/GARCH) models [11], and (h) count data forecasting [12].

In reservoir engineering, the prediction of hydrocarbon or water rates from geological formations is a time-series forecasting problem, with empirical solutions developed by Arps [13], referred to as decline curve analysis (DCA) technique. The method is based on a curve-fit principle, where one would attempt to fit either exponential, hyperbolic or harmonic curve to historical flow production rate as a function of time. Equation 2 shows the general form of the equation

|  |  |  |
| --- | --- | --- |
|  |  | Equation 2 |

where qi is the initial rate (bbls/day), Di is the initial decline rate (units) and b is the degree of curvature of the line. An exponential curve fit would have b = 0, a hyperbolic curve would have 0 < b < 1 and a harmonic curve would have b = 1. The fitted curve is then used to predict future production rates and cumulative production. This method was originally designed to work with high porosity-permeability reservoirs, and tends to overestimate hydrocarbon recovery from unconventional (low permeability) reservoirs. Thus, various authors have tried to expand on this work [14, 15, 16, 17], but for all intends and purposes, they are mostly just variations of the initial DCA method developed by Arps. For an effective DCA forecast, domain knowledge is key, but inherently the process is one of trial and error, and thus it is not uncommon for DCA results to have low-best-high estimates.

With the advent of big data, fast computing and cheap memory, applying a machine learning (ML) and artificial intelligence (AI) solution for time-series forecasting seems a natural evolution. ML solutions were first introduced to the petroleum industry in the early 2000s. Application of ML and AI include addressing prediction of reservoir parameters [18], history matching, of oil, gas and water production forecasting (flow rate prediction), pattern recognition in well logs and well tests analysis, production enhancement and prediction of failures, among others [19, 20, 21].

However, in the domain of CCUS and injector production performance, there is little to no available data related to time series and forecasting. Rather, literature appears to be mostly concentrated in predicting carbon emissions [22], leakage [23], CO2 absorption and adsorption [24], property prediction and process simulation [25], simulation of transportation, and geological behaviours as it relates to uncertainty analysis, sequestration, utilisation and EOR processes [26, 27, 28]. Work by Iskander et al perhaps most closely approaches this, where Long Short-Term Memory (LSTM) networks were utilised to forecast oil, water and CO2 production at future infill well locations, for both single phase and 3-phase fluid models. Data was in the form of a synthetic PUNQ-S3 reservoir model, combined with real-world observations from 8 production wells, which recorded daily production volumes over a decade from 2004-2014 [29].

We aim to develop on the work of Iskander and others by utilising ML and AI methods, and in particular LSTM, but focusing on the prediction of CCUS injection well performance, using the open-source information from the IBDP. We will demonstrate how our developed LSTM model shows a correlation between the change in injection rate to the behaviour of other dynamic parameters [29].

While the primary purpose of the model is as a checkpoint against carbon migration, either at the well location or from other losses during the injection process, we view the model as another means for engineers to perform scenario based de-risking of exploration plays, via modelling variation in well and storage parameters to validate CO2 containment. The model will also aid in the understanding of the injection process and potentially can be used to “right size” well operations and optimise costs.

# Scope and Methods

## Datasets

During this three-year period, a substantial amount of data was collected from both an injection and monitoring well, 700 ft apart. The injection well was drilled to a total depth (TD) of ~7051 ft, and was drilled with a 26” bit to 355 ft, and cased with a 20” casing to surface. A 17 ½” hole size than followed to a TD of 5339 ft, and an intermediate casing string 13 3/8” in diameter was set. The reservoir section was drilled in a 12 ½” hole size to ~ 7056 ft, and competed with a 9 5/8”production casing and 4 ½” tubing. The perforations were made at the base (i.e. above the pre-Cambrian) Mt. St Simon Sandstone, which was a relatively thick reservoir of ~1620 ft. A total of 3 geophones were set at 4925 ft, 5743 ft and 6137 ft as well as a pressure / temperature gauge mandral at 6325 ft. The monitor well was drilled to a total TD of 7272 ft; it had a surface casing (13 3/8” to 377 ft, followed by intermediate casing of 9 5/8” to 5322 ft and 5 ½” casing across the Mt. St Simon Sandstone, which contained a 3 component geophone array [30].

1. *Data Collection and Preparation*

A total of 34 parameters were measured from the injection and verification well. The parameters measured comprised of both surface and downhole measurements. The time-scale of the data is over three full years, at hourly intervals (27,464 rows/hours). 27,263 hours of data will be used to train a suitable model which will be use to predict 201 hours of D in the future. Given in **Table 1** is the descriptive statistics of the provided data. We note that there are a series of “null values” and non-nuemric numbers, and also note that for some of the input data measurements, there are significant outliers

**Table 1:** Data Statistics

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Measurement | Non-Zero  Value | Mean Value | Standard Deviation | Minimum Value | 25th  Percentile | Median Value | 75th  Percentile | Maximum Value |
| Avg\_PLT\_CO2VentRate\_TPH | 27398.0 | 2.1 | 133.2 | 0.0 | 0.0 | 0.1 | 0.2 | 18333.2 |
| Avg\_CCS1\_WHCO2InjPs\_psi | 27270.0 | 1239.9 | 817.7 | 0.0 | 1235.5 | 1338.9 | 1361.0 | 39032.4 |
| Avg\_CCS1\_WHCO2InjTp\_F | 27398.0 | 89.8 | 48.3 | 0.0 | 93.0 | 96.3 | 96.9 | 2879.4 |
| Avg\_CCS1\_ANPs\_psi | 27304.0 | 560.9 | 445.9 | 0.0 | 523.5 | 564.9 | 604.8 | 24105.6 |
| Avg\_CCS1\_DH6325Ps\_psi | 27398.0 | 3244.2 | 173.5 | 0.0 | 3233.0 | 3286.1 | 3324.7 | 3515.9 |
| Avg\_CCS1\_DH6325Tp\_F | 27398.0 | 127.7 | 7.2 | 0.0 | 127.2 | 130.1 | 131.1 | 135.7 |
| Avg\_VW1\_WBTbgPs\_psi | 26127.0 | 1801.8 | 999.4 | 0.0 | 2173.5 | 2322.4 | 2379.8 | 4954.7 |
| Avg\_VW1\_WBTbgTp\_F | 26061.0 | 80.8 | 44.3 | 0.0 | 103.4 | 104.2 | 105.0 | 120.1 |
| Avg\_VW1\_ANPs\_psi | 23487.0 | 525.0 | 3988.7 | 0.0 | 0.5 | 4.7 | 16.9 | 31993.5 |
| Avg\_VW1\_Z11D4917Ps\_psi | 26688.0 | 1597.5 | 873.3 | 0.0 | 2070.3 | 2073.4 | 2074.1 | 2378.0 |
| Avg\_VW1\_Z11D4917Tp\_F | 26709.0 | 81.7 | 44.2 | 0.0 | 103.7 | 105.2 | 106.5 | 108.6 |
| Avg\_VW1\_Z10D5001Ps\_psi | 26688.0 | 1627.9 | 889.9 | 0.0 | 2106.9 | 2112.4 | 2116.6 | 2420.6 |
| Avg\_VW1\_Z10D5001Tp\_F | 26709.0 | 81.2 | 44.0 | 0.0 | 101.8 | 104.7 | 105.0 | 110.9 |
| Avg\_VW1\_Z09D5653Ps\_psi | 26688.0 | 1961.1 | 1071.9 | 0.0 | 2534.0 | 2547.5 | 2551.1 | 2785.4 |
| Avg\_VW1\_Z09D5653Tp\_F | 26709.0 | 87.4 | 47.3 | 0.0 | 111.5 | 112.8 | 113.5 | 114.9 |
| Avg\_VW1\_Z08D5840Ps\_psi | 26189.0 | 1604.8 | 1289.3 | 0.0 | 0.0 | 2627.3 | 2637.4 | 4446.2 |
| Avg\_VW1\_Z08D5840Tp\_F | 25878.0 | 69.0 | 56.4 | 0.0 | 0.0 | 114.0 | 115.0 | 353.2 |
| Avg\_VW1\_Z07D6416Ps\_psi | 25985.0 | 2199.2 | 1273.0 | 0.0 | 0.0 | 2911.6 | 2925.3 | 3195.0 |
| Avg\_VW1\_Z07D6416Tp\_F | 25985.0 | 88.8 | 50.8 | 0.0 | 116.5 | 116.9 | 118.6 | 145.1 |
| Avg\_VW1\_Z06D6632Ps\_psi | 25500.0 | 2315.6 | 1308.6 | 0.0 | 3012.2 | 3026.1 | 3031.3 | 3380.4 |
| Avg\_VW1\_Z06D6632Tp\_F | 25500.0 | 90.9 | 50.8 | 0.0 | 116.6 | 118.6 | 119.4 | 124.3 |
| Avg\_VW1\_Z05D6720Ps\_psi | 23955.0 | 2106.0 | 1447.2 | 0.0 | 0.0 | 3069.6 | 3073.6 | 3365.9 |
| Avg\_VW1\_Z05D6720Tp\_F | 23955.0 | 82.3 | 55.0 | 0.0 | 0.0 | 118.5 | 119.4 | 122.1 |
| Avg\_VW1\_Z04D6837Ps\_psi | 26600.0 | 2347.5 | 1369.0 | 0.0 | 0.0 | 3148.0 | 3153.5 | 3331.9 |
| Avg\_VW1\_Z04D6837Tp\_F | 26600.0 | 91.0 | 51.6 | 0.0 | 118.8 | 119.5 | 119.9 | 125.8 |
| Avg\_VW1\_Z03D6945Ps\_psi | 24361.0 | 2350.9 | 1495.7 | 0.0 | 0.0 | 3299.9 | 3320.6 | 3457.9 |
| Avg\_VW1\_Z03D6945Tp\_F | 25932.0 | 182.3 | 368.5 | 0.0 | 0.0 | 121.3 | 122.8 | 1602.9 |
| Avg\_VW1\_Z02D6982Ps\_psi | 26423.0 | 2456.6 | 1459.1 | 0.0 | 0.0 | 3316.2 | 3332.0 | 3499.6 |
| Avg\_VW1\_Z02D6982Tp\_F | 26423.0 | 91.3 | 52.7 | 0.0 | 32.0 | 121.3 | 122.0 | 124.3 |
| Avg\_VW1\_Z01D7061Ps\_psi | 25307.0 | 2300.2 | 1521.4 | 0.0 | 0.0 | 3318.2 | 3327.3 | 3445.1 |
| Avg\_VW1\_Z01D7061Tp\_F | 25108.0 | 85.9 | 55.4 | 0.0 | 0.0 | 121.4 | 122.6 | 133.9 |
| Avg\_VW1\_Z0910D5482Ps\_psi | 26709.0 | 1855.6 | 1017.6 | 0.0 | 2353.3 | 2374.9 | 2416.1 | 2758.3 |
| Avg\_VW1\_Z0910D5482Tp\_F | 26709.0 | 86.3 | 46.7 | 0.0 | 110.5 | 111.5 | 112.0 | 113.7 |
| inj\_diff | 27397.0 | 0.0 | 82.7 | -11021.1 | -0.1 | 0.0 | 0.1 | 7033.5 |

1. *Cleaning Data*

We reviewed the attributes for the data, and were especially concerned about discontinues and non-numerial data. We utilised 3 methods to clean the data, in order of operation (i) we firstly performed a computational fill of all the missing and null values with a “forward fill’ operation, with the last valid observation being propagated forward. We do this on the assumption that missing values retain the properties of the previous cell i.e. there has been no change in the data between time *t* and time (*t-1*), (ii) a Z-score method where

|  |  |  |
| --- | --- | --- |
|  |  | Equation 2 |

xi is a single data value, is the median of the data set and MAD is the median absolute deviation of the dataset and (iii) a visual check of the data, removing outliers that we view as being deleterious to the interpretation. Shown in Figure 2 is an example of the impact that data cleaning has on the data quality. We see that our process has removed spikiness in the data and smoothed out some of the small scaled perturbations, resulting in a more manageable dynamic range.

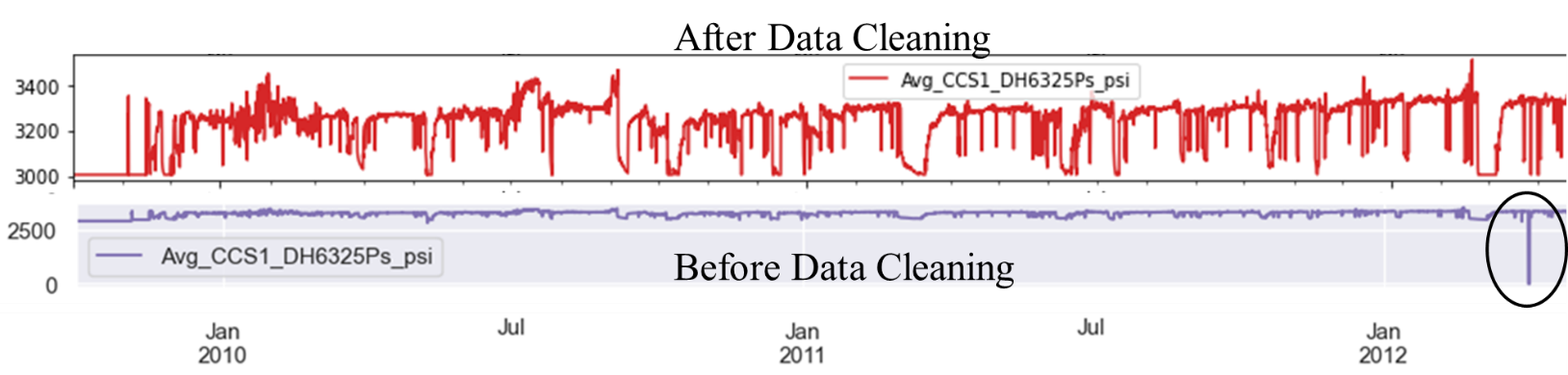


Figure 2: An example of the measurement before (lower image) and after (upper image) data cleaning

1. *Machine Leaning Model Selection - Long Short-Term Memory (LSTM) vs. Autoregressive Integrated Moving Average (ARIMA)*

In the choice of ML model to apply, we had to decide between a statistical approach based on ARIMA, or alternatively, the use of a non-linear algorithm such as neural networks (NN). Many authors have looked into comparing such methods in various fields. What we noted from our review of the work was there was a general agreement that approaches like ARIMA would not only require less inputs, but would be less of a “black-box”, which NNs are known to be [30, 31, 32].

ARIMA has been applied by previous authors to forecast oil production data [33]. The “autoregressive” piece of ARIMA deals with finding a correlation between a specific value and a prior/lagged value, essentially seeing if a variable has any correlation to its past values. The “integrated” piece deals with making data stationary, essentially ensuring that properties of the data (such as mean and variance), are constant over time. The “moving average” piece of the model finds the dependency between a specific value, and the error from a moving average model applied to previous values. ARIMA models are therefore useful in forecasting time series data and are especially useful when trying to predict time series data that is non-stationary. While Ning et al observed that ARIMA was robust in predicting rates of oil production across wells, our review of ARIMA models being produced with high frequency data, where accuracy on an hourly basis was important, found that the error rate compounds significantly when the forecasting horizon is extended beyond a day [34].

LSTM is a type of RNN which uses useful patterns from sequential data to provide accurate forecasts [19]. It learns from previous outputs to provide better results the following time. A typical LSTM has 3 layers (i) an input gate which assigns weights based on the significance of different variables, (ii) a forget gate to retain only useful information, and (iii) an output gate which manages the information flow. LSTM holds a memory cell which retains captured information over longer time periods and preserves useful constituents using its input and forget gates, hence avoiding the vanishing gradient issue associated with traditional NNs. LSTMs are particularly useful for non-linear problems where there does not appear to be strict mathematical relationships between variables. In our review of LSTM vs ARIMA models, we have found a common consensus in various fields that LSTM performed better, with reduced error rates but with significantly increased processing time [32, 31].

|  |
| --- |
| Comprehensive guide to LSTM & RNNs.  Introduction to LSTM Units in RNN | Pluralsight |

Figure 2: Schematic of LSTM network

We settled on the choice of LSTM because we realised that the data we were analysing was likely to contain non-seasonal, high frequency information, where accuracy was going to be important [35]. Additionally, the were numerous variables provided which we did not know to be important in prediction without the model build.

1. *Model Training and Hyperparameter Optimization*

We initially utilized 20 different algorithms but finally opted for 3 based on the ranked results of the mean absolute error (MAE), the mean squared error (MSE), the root mean squared error (RMSE) and the coefficient of determination (R2):

where yi is the predicted value, xi is the true value, is the mean of xi and n is the total number of data points. Note that a good result is one that has a low MAE, MSE and RMSE, but a high R2 (since this is essentially a measure of “goodness of fit”).

The models were then tuned by varying “hyperparameters”. A hyperparameter is a characteristic of a model that is external to the model and whose value cannot be estimated from data. The hyperparameters are optimized via a search algorithm with the goal of minimize the overall error metric.

To account for idiosyncrasies in the data (noise, patterns, outliers, etc.), k-fold cross-validation was run to validate the stability of the model.

1. *Testing, Deployment, and Evaluation of model against our blind datasets*

# Results and Discussion

|  |  |
| --- | --- |
|  |  |

Figure 3: Correlation Matrix – Pairwise Correlation

We observed that for the TORIS database, “lithology code” and “depositional system” show collinearity, as did the “Temperature-Pressure” pair. For the GOM database, “Pressure-Depth” and “Gas Oil Ratio-Formation Volume Factor” were collinear. The relationship between these variable pairs aligns with conventional reservoir engineering concepts. As keeping both parameters will not add additional information to the predictive model and may potentially result in an overfit within the model, a single element from each of the variable pairs is eliminated to allow for a more stable model.

The final input variables were determined and summarized in Table 1 for the TORIS dataset and Table 2 for the GOM dataset.

The TORIS dataset comprises 10 categorical data types, and 14 numerical data types. The total data size is 389 values per column, for a total of 9,336 data points which was approximately ~9.6% of the original data base.

Table 2: Statistics of the TORIS Database Input Variables

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | count | mean | std | min | 0.25 percentile | 0.5 percentile | 0.75 percentile | max |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

*Initial Machine Leaning Model*

*Tuning and Optimizing the Model*

*Evaluation against the “Hold-out” Data and the “Double Blind” Set*

# Discussion & Recommendations

Below we summarise our key findings and learnings in this of developing the ML model for RF prediction.

*Data Preparation*

When performing the initial data-analysis, it is essential to combine both data analysis techniques with domain knowledge. In both datasets, many values are missing. This can be dealt with by dropping missing variables. However, a better alternative would be to use domain knowledge to try to supplement the missing data. The use of domain knowledge and experience is critical because correlation and causality are not the same. Taking the example of “STOIIP vs Porosity” and “Gross Pay vs Porosity”, a correlation relationship is geologically reasonable in the former, but not in the latter.

*Optimization*

In an optimalization problem, the end goal is the minimization of a metric/ cost function. In our case, as we are dealing with regressors, the minimization must be applied to error functions like MAE, MSE and RMSE. The solution is best achieved via a gradient descent algorithm where the minimum of the cost function is achieved iteratively, largely because as the negative of the gradient is followed over time, it would theoretically reach a point where it will no longer be possible to decrease the cost function any further. In other words, when the number of iterations increases, the solution moves towards the minima which is defined by an optimal input hyperparameter set. The challenge is than finding that sweet spot of iterations and hyperparameters, and accepting the tradeoffs that occur, especially as it related to training time.

*Parametric Feature Importance*

When looking at the ranked feature importance, we observed that categoric data was not considered to be critical in any of the selected ML algorithms. The effects of categoric data, like trapping type, lithology or diagenetic overprint might be important characteristics in certain reservoirs. We think that the underrepresentation is due to the plethora of possible individual values. For example, in TORIS, there are 54 unique depositional environments (eolian, lacustrine, shelf, reefs, pinnacles etc). To the ML model, there are insufficient training instance of each to observe a strong trending behaviour. To increase the importance of categoric data, the categories should be grouped and simplified. A recommendation is for a skilled geologist or geophysicist with some background in data science to look at helping to simplify some of the categoric input parameters such that only 3 to 5 unique instances present.

*Conventional vs ML*

The developed ML models outperformed conventional empirical correlations such as Arps et al and Guthrie and Greenberger. The spread of datapoints was much wider with using both the Arps et al. (R2=0.21 and 0.006 in the TORIS and GOM model respectively) and Guthrie and Greenberger correlations (R2=0.15 and 0.007 in the TORIS and GOM model respectively). Further work would be to investigate the machine learning model and its comparative performance on other fields.

*Non-linearity*

Complementary to this work is the use of Artificial Neural Networks (ANN), which are a type of deep learning network complementary to ML. The key benefit of ANN is the ability to better handle non-linearities. Future work may investigate developing ANN networks to study the criticality of domain expertise and just how different RF predictions are if ANN is utilized instead of just ML.

*Sensitivities*

# Conclusion

In this work, we have demonstrated that ML models form a good basis for estimating RF; however, applying general domain knowledge and sense-checking results is still very important. The use of any ML model is dependent on the purpose of the RF estimation and should be complementary (rather than contrasting) to conventional techniques (whether they be using analogs or empirical methods). Any ML model should not replace the need for decline based, or simulation-based estimates in fields with extensive production history. In such instances, ML can be used to determine ultimate RF potential. For early-stage RF estimate, ML models might perform better that Arps et al. correlation and Guthrie and Greenberger correlation, especially when data is sparse.

# Reference

|  |  |
| --- | --- |
| [1] | R. Lindsey, “Climate Change: Atmospheric carbon dioxide, NOAA Climate.gov,” [Online]. Available: https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide. [Accessed 25 April 2023]. |
| [2] | United States Environmental Protection Agency, “Sources of Greenhouse Gas Emissions,” 20 Jul 2022. [Online]. Available: from https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions. |
| [3] | J. G. D. Gooijer and R. J. Hyndman, “25 years of time series forecasting,” *International Journal of Forecasting,* vol. 22, no. 3, pp. 443-473, 2006. |
| [4] | E. S. Gardner, “Exponential smoothing: The state of the art,” *Journal of Forecasting,* vol. 4, no. 1, pp. 1-28, 1985. |
| [5] | R. D. Snyder, “Recursive Estimation of Dynamic Linear Models,” *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 47, no. 2, p. 272–276, 1985. |
| [6] | R. H. Scumway and D. S. Stoffer, Time Series Regression and ARIMA Models, New York: Springer, 2000. |
| [7] | G. Huyot, K. Chiu, J. Higginson and N. Gait, “Analysis of Revisions in the Seasonal Adjustment of Data Using X-11-Arima Model-Based filters,” *International Journal of Forecasting,* vol. 2, no. 2, pp. 217-229, 1986. |
| [8] | S. H. Babbs and K. B. Nowman, “Kalman filtering of generalized Vasicek term structure models,” *Journal of financial and quantitative analysis,* vol. 34, no. 1, pp. 115-130, 1999. |
| [9] | M. P. F. P. H. &. S. N. R. Clements, “Forecasting economic and financial time-series with non-linear models,” *International journal of forecasting,* vol. 20, no. 2, pp. 169-183, 2004. |
| [10] | B. K. Ray, “Long-range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model,” *International Journal of Forecasting,* vol. 9, no. 2, pp. 255-269, 1993. |
| [11] | S. J. Taylor, “Forecasting the volatility of currency exchange rates,” *International Journal of Forecasting,* vol. 3, no. 1, pp. 159-170, 1987. |
| [12] | T. R. Willemain, C. N. Smart, J. H. Shockor and P. A. DeSautels, “Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method,” *International Journal of Forecasting,* vol. 10, no. 4, pp. 529-538, 1994. |
| [13] | J. J. Arps, “Analysis of Decline Curves,” *Transactions of the AIME,* Vols. SPE-945228-G, 1945. |
| [14] | A. L. L. a. T. P. Clark, “Production Forecasting with Logistic Growth Models,” in *SPE Annual Technical Conference and Exhibition*, Denver, Colorado, USA, 2011. |
| [15] | A. Duong, “An Unconventional Rate Decline Approach for Tight and Fracture-Dominated Gas Wells,” in *Canadian Unconventional Resources and International Petroleum Conference.* , Calgary, Alberta, Canada, 2010. |
| [16] | D. R. J. A. P. A. D. a. T. A. B. Ilk, “Exponential vs. Hyperbolic Decline in Tight Gas Sands — Understanding the Origin and Implications for Reserve Estimates Using Arps' Decline Curves.,” in *Paper presented at the SPE Annual Technical Conference and Exhibition*, Denver, Colorado, USA, 2008. |
| [17] | P. Valko, “Assigning value to stimulation in the Barnett Shale: a simultaneous analysis of 7000 plus production hystories and well completion records,” in *SPE Hydraulic Fracturing Technology Conference* , Woodlands, Texas. , 2009. |
| [18] | M. Kumar, K. Swaminathan , A. Rusli and A. Thomas-Hy, “Applying Data Analytics & Machine Learning Methods for Recovery Factor Prediction and Uncertainty Modelling,” *SPE Asia Pacific Oil & Gas Conference and Exhibition,* Vols. SPE-210769-MS , pp. 1-12, 2022. |
| [19] | N. Ibrahim, A. Alharbi, T. Alzahrani, A. Abdulkarim, I. Alessa, A. Hameed, A. Albabtain, D. Alqahtani, M. Alsawwaf and A. Almuqhim, “Well Performance Classiﬁcation and Prediction: Deep Learning and Machine Learning Long Term Regression Experiments on Oil, Gas, and Water Production,” *Sensors,* vol. 22, no. 5326, pp. 1-22, 2022. |
| [20] | P. Panja, W. Jia and B. Mcpherson, “Prediction of well performance in SACROC field using stacked Long Short-Term Memory (LSTM) network,” *Expert Systems With Applications ,* vol. 205, no. 117670, pp. 1-25, 2022. |
| [21] | Y. Dong, Y. Zhang and X. Cheng, “Reservoir Production Prediction Model Based on a Stacked LSTM Network and Transfer Learning,” *ACS Omega,* vol. 6, pp. 34700-34711, 2021. |
| [22] | J. Qin and N. Gong, “The estimation of the carbon dioxide emission and driving factors in China based on machine learning methods,” *Sustainable Production and Consumption,* vol. 33, pp. 218-229, 2022. |
| [23] | F. Hussin, S. A. N. M. Rahim, N. S. M. Hatta, M. K. Aroua and S. A. Mazari, “A systematic review of machine learning approaches in carbon capture applications,” *Journal of CO2 Utilization,* vol. 71, pp. 102474 - 102490, 2023. |
| [24] | M. Rahimi, S. M. Moosavi, B. Smit and T. A. Hatton, “Toward smart carbon capture with machine learning,” *Cell Reports Physical Science Perspective,* vol. 2, no. 100396, pp. 1-19, 2021. |
| [25] | Y. Yan, T. N. Borhani , S. G. Subraveti , K. N. Pai, V. Prasad, A. Rajendran , P. Nkulikiyinka , J. O. Asibor , Z. Zhang , D. Shao , L. Wang , L. Zhang , Y. Yong, W. Ampomah , J. You , M. Wang , E. J. Anthony , V. Manovic and P. T. Clough , “Harnessing the power of machine learning for carbon capture, utilisation, and storage (CCUS) – a state-of-the-art review,” *Energy Environ. Sci.,,* vol. 14, pp. 6122-6157, 2021. |
| [26] | P. Yao, Z. Yu, Y. Zhang and T. Xu, “Application of machine learning in carbon capture and storage: An in-depth insight from the perspective of geoscience,” *Fuel,* vol. 333, no. 1, p. 126296, 2023. |
| [27] | B. Chen, D. R. Harp, Y. Lin, E. H. Keating and R. J. Pawar, “Geologic CO2 sequestration monitoring design: A machine learning and uncertainty quantification based approach,” *Applied Energy,* vol. 225, no. 1, pp. 332-345, 2018. |
| [28] | J. You, W. Ampomah, Q. Sun, E. J. Kutsienyo, R. S. Balch, Z. Dai, M. Cather and X. Zhang, “Machine learning based co-optimization of carbon dioxide sequestration and oil recovery in CO2-EOR project,” *Journal of Cleaner Production,* vol. 260, p. 120866, 2020. |
| [29] | U. P. Iskander and M. Kurihara, “Long Short-term Memory (LSTM) Networks for Forecasting Reservoir Performances in Carbon Capture, Utilisation, and Storage (CCUS) Operations,” *Scientific Contributions Oil & Gas,* vol. 45, no. 1, pp. 35 - 50, 2022. |
| [30] | M. Alsharif, M. Younes and J. Kim, “Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea.,” *Symmetry,* vol. 11, no. 2, p. 240, 2019. |
| [31] | S. Siami-Namini, N. Tavakoli and A. S. Namin, “A comparison of ARIMA and LSTM in forecasting time series,” *17th IEEE international conference on machine learning and applications (ICMLA),* pp. 1394-1401, 2018. |
| [32] | S. A. Ludwig, “Comparison of time series approaches applied to greenhouse gas analysis: ANFIS, RNN, and LSTM.,” *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE),* pp. 1-6, 2019. |
| [33] | Y. Ning, H. Kazemi and P. Tahmasebi, “A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet,” *Computers & Geosciences,* vol. 164, no. 105126, 2022. |
| [34] | A. Parasyris, G. Alexandrakis and G. Kozyrakis, “Predicting Meteorological Variables on Local Level with SARIMA, LSTM and Hybrid Techniques,” *Atmosphere,* vol. 13, no. 878, 2022. |
| [35] | G. Reikard, “Predicting solar radiation at high resolutions: A comparison of time series forecasts,” *Solar Energy,* vol. 83, no. 3, p. 342–349, 2009. |

# Acknowledgements