Groundwater Level Forecasting: USA

MSc Research Project
MSCDAD

Adeola Deborah Adeniji
Student ID: X23104201

School of Computing
National College of Ireland

Supervisor: Dr David Hamill

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Adeola Deborah Adeniji
Student ID:	X23104201
Programme:	MSc. Data Analytics **Year:** 2024
Module:	MSc. Research Project
Supervisor:	Dr David Hamill
Submission Due Date:	12/12/2024
Project Title:	Groundwater Level Forecasting: USA
Word Count:	7234
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Adeola Deborah Adeniji

Date: 11th December 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Groundwater Level Forecasting: USA

Adeola Deborah Adeniji
X23104201

Abstract

Beneath the earth's surface lies a natural reserve, groundwater, a vital source for drinking water, agriculture, and ecosystem sustainability in arid areas like California, USA. Groundwater management is of utmost importance but faces challenges due to its fluctuation levels and uncertain supplies. With the urgent need to ensure the rational use of this reserve, the Department of Water Resource (DWR), needs to provide a reliable and accurate forecast for groundwater levels, to support California water resource management (CDWR). This research leverages historical data of over 30 years to evaluate the effectiveness of time series models – Simple Time Series, Exponential Smoothing and ARIMA Models, to forecast groundwater levels for the next 8 years. Multiple performance metrics were used to find the best model for long-term water management, including coefficient of determination (R^2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Akaike's Information Criterion (AIC), and Mean Absolute Percentage Error (MAPE). The results indicated that ARIMA model outperformed the others, achieving the lowest RMSE of 2.31 and the highest R^2 of -0.04. However, due to the complex dynamics of groundwater fluctuations, all models including Arima, struggled and were unable to provide satisfactory accuracy for the forecast. These findings underline the need to adopt more advanced methods since current models did not achieve good accuracy that can provide decision-makers useful information, to support sustainable groundwater management. This research contributes to the knowledge base, by identifying the models' limitations while suggesting alternative ways of improving groundwater level forecasting.

Keywords: Groundwater levels forecasting, time series models, department of water resources (USA), ARIMA prediction, sustainability groundwater use.

1 Introduction of Groundwater Level Forecasting

Groundwater (GW), a life vital natural reserve found beneath the Earth's surface, is stored within spaces between soil particles and rocks (Lall et al., 2020). It accounts for greater than 30% of the world's freshwater supply and is a vital source for drinking water, agriculture, and sustainability within ecosystems. GW is stored and filtered by soil, sand, and rock as it surfaces from underground springs or wells, this makes it less vulnerable to contamination, more sustainable and a reliable resource, unlike surface water which is directly exposed to pollution. In California, for instance, during dry years, up to 60% of their water supplies comes from GW, as this reveals its importance, especially in the arid regions (California DWR, 2024). While GW management is of utmost necessity, it faces challenges due to the uncertainties in water supply, low annual rainfall and fluctuations of aquifers levels. This complexity makes planning for effective resource management difficult. To better understand these challenges, continuous monitoring of groundwater levels (GWL) is needed to make informed decisions for sustainable water use.

This study is motivated by an urgent need that has become completely unavoidable to develop an efficient model that can accurately predict GWL. The capability of these predictions is important for tracking and balancing the supply and demand, as well as creating sustainable water resource management practices. This research leverages the use of data mining methods for time series extraction from GW measurement data, with the aim of addressing the current challenges in the water supply. By integrating a well-fitted predictive model into a monitoring system, this approach will provide predictions that facilitate timely decision-making.

Furthermore, section 2 reviews numerous studies that have predicted groundwater levels using various data-driven approaches. These include deep learning models like Artificial Neural Network (ANN), machine learning models like Support Vector Machine (SVM), and time series models like ARIMA. While ANN and SVM are effective models, their "black box" nature makes the interpretation of their results difficult for decision-makers (Rajaee et al. 2019**;** Boo et al., 2024). On the other hand, the ARIMA model has shown better interpretability in GWL predictions, which makes it easy for decision-makers (Sarma et al., 2022**;** Takafuji et al., 2019). While indicating the need to investigate a more effective model for reliable forecasting, most existing studies focus on short-term data range and predictions, which limits their applicability for long-term resource management. This study proposes addressing these limitations, with a model that can forecast GWL for up to eight years, using over 30 years of historical data. This has led to the research question below of this study.

1.1 Research Question

To what extent can time series models forecast groundwater levels to support sustainable water resource management in the United States?

1.2 Research Objectives for Groundwater Level Forecasting

To address the research question, key objectives were accomplished. The first objective conducted extensive literature reviews, second objective collected, analyzed and prepared GWL data. The third objective implemented statistical tests to validate parameters. The fourth objective developed time series models including Simple Time Series, Exponential Smoothing and ARIMA models with descriptive visualizations. The fifth objectives evaluated the performance of all models using metrics like R^2, RMSE, MAE, AIC, and MAPE. Finally, the study compared models based on R^2 and RMSE performance.

1.3 Structure of Research

This paper is organized as follows: Section 1 provides introductory information on the research study. Section 2 investigates previous related work on GWL forecasting. Section 3 presents methods specifically to achieve the objective. Section 4 presents the implementation and experimental setup of the research. Section 5 evaluates and compares the results of the models. Finally, Section 6 concludes by discussing the results and future work that are applicable in answering the research question in section 1.1.

2 Related Work on Groundwater Forecasting

In this section, related works conducted over the last five years on GW forecasting will be discussed, as it highlights an overview of its key challenges, different data-driven approaches, their comparisons, gaps identified in the literature and how this study addresses them. It finally concludes on how these insights answer the question in the previous subsection 1.1.

2.1 Overview of Groundwater Management Challenges

The importance of forecasting groundwater has long been investigated. A recent review paper analysed over 168 articles that was published on GWL forecasting between 2000 to 2023 (José Luis Uc-Castillo et al., 2023). The study reveals that most research was done in semi-arid or arid regions like the United States where surface water is scarce and unreliable. Another systematic review with over 109 research articles from 2008 to 2002 further supported these views, where surface water is both limited and highly valuable (Khan et al., 2023). These reviews underlined the importance of the role GW plays in the area for sustainable water resource management, particularly in ensuring drought resilience, meeting in the increased demand of water supply, supporting ecosystem health, and maintaining water quality and economic stability. These studies further emphasise the growing need as to why enhanced accuracy is important in predicting GWL (Paul et al., 2024).

However, GW management sustainability and effectiveness are threatened by several challenges stemming from data scarcity, model complexity, variability in results, and the need for accurate predictions. The lack of sufficient hydroclimate or GWL measurement data remains a barrier to integrating remote sensing and machine learning models (Abdellatif Rafik et al., 2023). The use of machine learning provides promising solutions as local models may achieve high accuracy, but this introduces another set of related challenges to the interpretability of the models (Shaikh et al., 2024). Aside from the technological challenges that groundwater management faces, some environmental factors are very vital to note, such as climate change, aquifer level fluctuation, and seasonal droughts (Aranguren-Díaz et al., 2024). Despite these challenges, the advancement in artificial intelligence and machine learning offers potential ways for improved groundwater management (Boo et al., 2024).

2.2 Data-Driven Approaches in Groundwater Level Forecasting

In recent years, many researchers have leveraged methodologies such as Machine Learning (ML), Deep Learning (DL) and Time series (TS) models on GW data (Sarma and Singh, 2022). The adaptation of these models has led to its own advantages and challenges in terms of performance, interpretability and sustainability in relation to long-term forecasting, as all these factors will be considered to answer the research question in section 1.1.

The predictive abilities of machine learning techniques were further tested in the context of multiscale GW forecasting, as it has the capability of modelling nonlinearities between GW and its environmental drivers such as rainfall (Rahman et al., 2020). With the lack of physical understanding, the machine learning approach in over two decades has been able to overlook the intricate hydrological processes as it relies on the statistical connection such as rainfall, linking with the response variables such as groundwater levels (Rajaee et al. 2019). An experimental exercise was conducted in the southern part of Japan where over one million people depended fully on groundwater for their daily domestic use with the usage of 58 pump stations to meet their demand. The exercise used three ML models; Support Vector Machine was used for its frequent accuracy predictions (Ferreira et al., 2019; Rajaee et al. 2019), Random Forest for its high precision in handling input variables that are large in numbers (Tyralis, Papacharalampous and Langousis, 2020), Extreme Gradient Boosting (XGBT) for its built-in improved regularization features that helps prevents overfitting which is a common problem for machine learning models (Chen et al., 2019). Even though machine learning has previously showed success in forecasting GWL, it exhibits some limitations such as inaccurate predictions when multiscale changes over time (Rahman et al., 2020). The study suggested that the result could further be improved with a more advance model that could identify hidden time frequencies from localised features.

The application of DL in GWL forecasting has gained more traction and has emerged as a powerful tool when predicting complex, and non-linear relationships with large datasets. A study was able to use Recurrent Neural Networks (RNNs) to predict GWL in a confined area across southern Africa (Seyler et al., 2020). In the research, the deep learning model was used to analyse the temporal behaviour, influx, reserves and discharges. With the goal to answer two questions, first, to what extent could deep learning further predict groundwater levels? and secondly, what efficiency is required to generate aquifer fluxes with the adapted deep learning model. Particularly, the study evaluated and compared the performance of two RNNs models including Neural Network Autoregressive (NNAR) and Long-Short Term Memory (LSTM) networks. With the use of LSTM as a variant of RNNs, it was found to be superior to NNAR in terms of its prediction accuracy, having better results on the performance metrics such as RMSE and R^2. In the realization of addressing the questions, the research highly noted LSTM to be efficient in identifying GWL changes and fluxes, while NNAR model struggled and was unstable to capture key variables in relation to the influx. In spite of the progress made, the research highlighted a significant limitation, that it lacked sufficient high-quality data which remains an important factor when using DL models for forecasting GWL (Seyler et al., 2020).

Capitalizing on the shortcomings of the previous research, another study extended the scope of its own research, as it applied different sets of machine and deep learning techniques using a more complex, high-quality dataset which aimed to predict groundwater levels in the United States across twelve districts (Bedi et al., 2020). The study showing its relatable properties to the objective of this research in section 1.1. The study included models like Support Vector Machine (SVM), Extreme Gradient Boosting (XGB), and Artificial Neural Networks (ANNs). From the data provided, the study focus was on the water quality, mainly considering the pesticides and nitrates levels, as the use of deep learning techniques proved to be insightful during the evaluation process. Based on the result, XGB had outperformed both SVM and ANN in terms of accuracy, however, a notable fault was found in the model. This highlights a questionable trade-off between practicality and accuracy when applying deep learning techniques to environmental data

Subsequently, another researcher used over 20 years of historical data on GWL to make predictions for the next five years in Taiwan. The study used both LSTM and convolutional Neural Networks (CNNs) as the deep learning techniques, after preprocessing and dividing the data into training, validation, and test sets (Chen et al., 2023). The models evaluated using standard metrics produced high accuracy values such as RMSE resulting in LSTM having 0.008, CNN with 0.007 and R^2 value of both models exceeding 0.997. From the results of the predictions, the deep learning model demonstrated that it had the ability to capture temporal dynamics of groundwater systems as it also identified some limitations (Chen et al., 2023).

2.3 Comparisons in Groundwater Forecasting Approaches

Numerous approaches have been investigated for forecasting GWL as each have shown their individual strengths and boundaries. For instance, a comprehensive review was conducted on over 109 papers, as it identified the gaps in both machine learning, deep learning approaches and traditional time series (Khan et al., 2023). During the investigation, ARIMA model stood out, as it performed better in capturing trends and seasonal fluctuations. Also, in light of the investigation, the study noted that most models were evaluated with common performance metrics including RMSE, MAPE, and R^2 as it underlines the importance of conducting long-term interval forecasting. This provided insights into the ARIMA model for further analysis.

While considering the traditional time series model, there have been some reoccurring deep learning models like LSTM and ANNs that have shown excellent results when capturing complex relationships, but they require extensive computational power (Seyler et al., 2020; Bedi et al., 2020). However, the deep learning models including LSTM and CNN that were used in predicting groundwater levels in Taiwan, were explained to have failed to uncover extreme patterns, which indicates a probable setback when dealing with high-variable data (Chen et al., 2023). In contrast, the use of the traditional time series models such as ARIMA when analysed by (Mirsanjari et al., 2019; Rashid et al., 2022) was discovered to have shown strengths in forecasting long-term trends, and seasonal variations.

When comparing the different data-driven approaches, certain realization shows that time Series models (particularly Arima), seem to be more reliable in forecasting trend-based fluctuations in GWL. The model has an inherent strength in handling temporal data and their simplicity (Takafuji et al., 2019). An example of its unique abilities over other approaches was where the ARIMA model was evaluated and compared with both ANNs and other statistical models in term of its accuracy, interpretability and simplicity (Kontopoulou et al., 2023). Regardless of the need for assumptions, the model was able to forecast without extensive input data, as it was less prone to overfitting which made it an appealing choice, especially for regions like Iran and Mahran with limited GW data (Goodarzi, 2020). While deep learning models offer flexibility, the ARIMA model which is proposed in this study, is been adopted due to its ability in handling long-term trend analysis and its interpretability which is essential for sustainable groundwater management (Takafuji et al., 2019).

2.4 Identifying the Gaps in Groundwater Levels Forecasting

Through the review of previous literature, GWL forecasting has faced some challenges resulting in gaps in current methodologies. Groundwater is a vital resource and of deep importance to certain regions, hence accurate forecasting of models is necessary for sustainable management (Lall et al., 2020). A summary table below has highlighted these limitations from a few research papers while suggesting a feasible solution that can improve the accuracy and applicability of groundwater forecasting models.

Table 1: Summary of Gaps and Solutions in Groundwater Level Forecasting

Reference	Gaps	Details	Solutions
Rashid et al.(2020); Seyler et al. (2020)	Short Dataset Duration	Many studies use only five to ten years of data, limiting long term insights.	Leverage over thirty years of historical data for a long-term trend
Rashid et al. (2022); Chen et al. (2023)	Short-Term Prediction	Most forecast often focus on one to three years predictions, missing short term sustainability needs	Extend prediction horizon to eight years for better long-term planning
Goodarzi (2020); Khan et al. (2023)	Complexity & Black-Box Nature of ML and DL Models	ML and DL models such as SVN and ANN are often difficult to interpret and explain	Use ARIMA model for simplicity, transparency and ease for interpretation
Bedi et al. (2020); Khan et al. (2023)	Limited Evaluation Metrics	Studies mostly use RMSE, MAE, and R^2, lacking comprehensive evaluation	Employ broader evaluation metrics such as R^2, RMSE, MAE, AIC, and MAPE for deeper analysis.

2.5 Conclusions of Related Work

In summary, analysing different research papers identified gaps in GW forecasting, and addressing them is important for an effective management. The use of a higher-quality dataset for a longer-term prediction on easier interpretable models with a wider range of evaluation metrics could further improve the reliability and relevance of this research. From the above literature, a foundation has been established for further investigation into the research propose, specifically adapting the time series models. The next section provides a detailed explanation on the methodology adapted for a novelty solution into forecasting groundwater as it aims to answering the research questions in section 1.1.

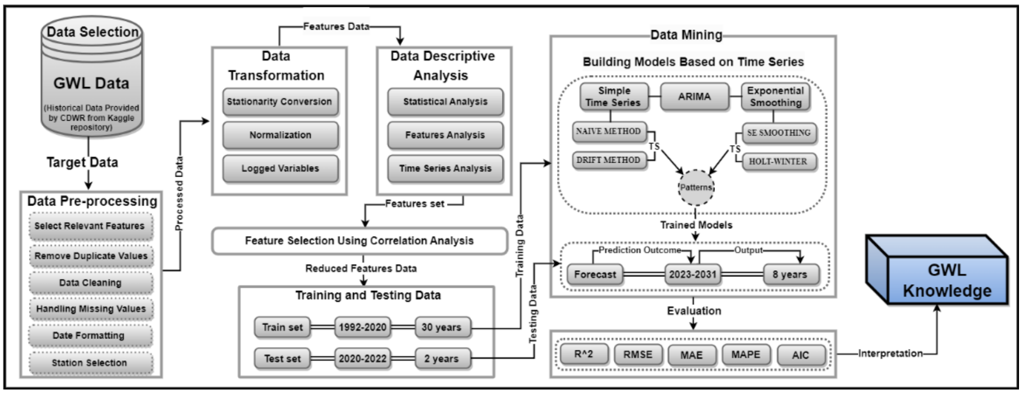
3 Research Methodology in Groundwater Level Forecasting

This section provides a description of the methodology used to forecast GWL. The methods propose an advanced and robust time series model for more sustainable water resource management. The subsection explains the adopted framework, architectural design, and process flow of the applied methods. Additionally, the subsections include diagrams and tables that better show each process activity.

3.1 Adopted Framework in Groundwater Levels Forecasting

In this research, the Knowledge Discovery in Databases (KDD) methodology is adopted as a framework in forecasting GWL. This is because KDD serves as a systematically structured and multi-phase process that can learn valuable information from historical data (Wetzel et al., 2024). The KDD framework shown in Figure 1. consists of eight major steps of a reproductive approach used to analyse time series data in other to build predictive models (Iqbal et al., 2021; Baydaroğlu et al., 2023). The functional flow of the diagram as used in this research consists of eight major steps. Step one, Data Selection, which involved collecting raw historical GW data from Kaggle repository provided by CDWR. Step two, Data pre-processing, where the raw data is been transformed and cleaned it to improve quality and efficiency. Step three, Data Transformation, refined the features to sure the data is in a suitable format for further analysis. Step four, Data Descriptive Analysis which explores underlying characteristics through time series and statistical analysis of the features. step five, which selects relevant features to optimize model accuracy. step six, Splits Data into training and testing set. In step seven, Model Building of various time series methods are developed using the training data, while the test data are used to evaluate the trained models. Finally, step eight, Evaluation, adopted multiple performance metrics to determine accuracy and reliability. The goal is to create an actionable groundwater knowledge that supports the objectives in section 1.2. The following subsection describes the core architecture of the research.

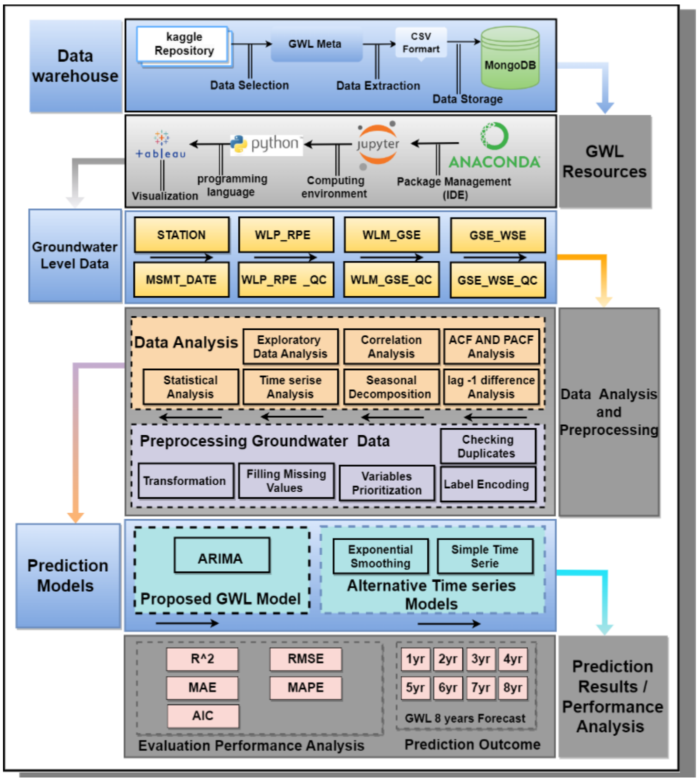
Figure 1: Adopted KDD Methodology for Groundwater Level Forecasting



3.2 Architectural Model Design in Groundwater Levels Forecasting

This subsection describes the core architecture used to develop the time series models. As illustrated in Figure.2, presents a comprehensive multi-layered architecture model design, that serves as a blueprint for GWL forecasting (Iqbal et al., 2021). The diagram comprises of six layers. The first layer begins with the data warehouse. At this stage, the data is extracted from Kaggle repository in csv format and then stored in a database (Meta data), to ensure easy access. The second layer shows the tools and software used for the research. These resources make the outlined objectives in section 1.2 achievable. The third layer shows the information of the data stored in layer-one and retrieved by layer-two. The fourth layer performs preprocessing steps on the GWL data from layer-two. While the preprocessing steps focused on transforming the GWL data into suitable formats for modelling. The fifth layer presents three different predictive models chosen for this research. First is the ARIMA model as the proposed and main model for the investigation. While the second models including Exponential smoothing and simple time series are alternative methods which are used as benchmarks to compare with the proposed model. The sixth layer is the final stage in the pipeline. At this layer, an evaluation in the performance of each model is analysed from the performance of the prediction results, and eight years forecast on different period is generated. Then the result evaluated provides knowledge about GWL prediction outcome in answering the research question. The next subsection provides a detailed process flow on how the architecture design was implemented.

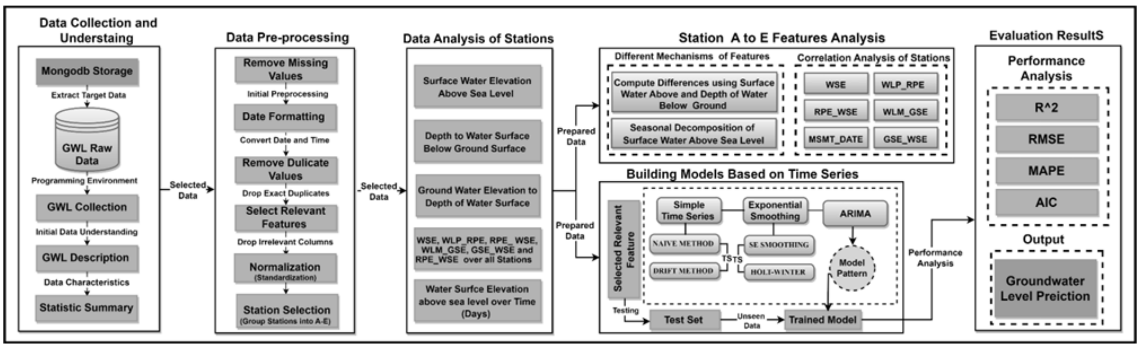
Figure 2: Architectural Model Design in Groundwater Levels Forecasting



3.3 Process Flow in Groundwater Levels Forecasting

A more detailed functional flow into each process method used in forecasting GWL is shown in Figure 3. The process begins by collecting and understanding the GWL raw data, extracted from the Kaggle repository and stored in a Mongo database. Next, the extracted data in the programming environment is then pre-processed, to a clean, organised and standard form. Once prepared, the selected data passes through multiple data analyses, where each feature across different stations is examined. The analysis showing the trends in patterns of the historical data supports the selection and application of building time-series models. The final evaluation produces output as GLW predictions. Next, subsection provides an overview of the GWL data[1] used in this research.

Figure 3: Process Flow Design in Groundwater Levels Forecasting



---

[1] https://www.kaggle.com/datasets/alifarahmandfar/continuous-groundwater-level-measurements-2023/data

3.4 Time-Series Groundwater Level Measurement Dataset

This subsection presents the GWL data used for this study, sourced from the California Department of Water Resources in Kaggle repository. Data titled "*Continuous Groundwater Level Measurement*" was made public in 2023 and license for usage by the open-source repository. The dataset consists of daily and monthly mean versions of GWL measurements, but in this study, the daily version of the data is utilized as it provides an extended range of information suitable for long-term predictions. The daily version of the data spanning from March 1992 to December 2020, contains over 1,048,575 records and 12 features of GWL information from the California region. Detailed properties of the GWL data are shown in table 2. The subsequent section provides information on how this data was further utilized.

Table 2: Detailed Summary of Continuous Groundwater Level Measurement Data

Column	Label	Data Type	Description
STATION	Station	Character (Chr)	Unique Station ID
MSMT_DATE	Measurement Date (PST)	Date (Date)	Date/time of measurement
WLM_RPE	RPE	Numeric (Float)	Reference Point Elevation
WLM_RPE_QC	RPE Quality Code	Integer (Int)	Quality code for WLM_RPE
WLM_GSE	GSE	Numeric (Float)	Ground Surface Elevation
WLM_GSE_QC	GSE Quality Code	Integer (Int)	Quality code for WLM_GSE
RPE_WSE	RPE to WSE	Numeric (Float)	Depth to water surface below RPE
RPE_WSE_QC	RPE to WSE Quality Code	Integer (Int)	Quality code for RPE_WSE
GSE_WSE	GSE to WSE	Numeric (Float)	Depth to water surface below GSE
GSE_WSE_QC	GSE to WSE Quality Code	Integer (Int)	Quality code for GSE_WSE
WSE	WS Elevation	Numeric (Float)	Water surface elevation above sea level
WSE_QC	WS Elevation Quality Code	Integer (Int)	Quality code for WSE

3.5 Data Understanding and Exploration

After collecting the GWL dataset, an initial exploration revealed potential issues. Notably, the acquired dataset had a total of 94,130 missing and duplicate values out of the 1,048,575 records as well as irrelevant features like the quality code columns shown in Table 2. that was not significant to the analysis. Also, inconsistent entities with the date column, made it not suitable for time series analysis, while the station column accumulated trivial location IDs, adding noise to the data. This finding guided the pre-processing and transformation step.

3.6 Data Pre-processing and Transformation

This step ensures that GWL raw data is properly cleaned and structured for a more consistent, quality and reliable forecasting. Missing and duplicate values from the data were removed, retaining over 954,445 records. The date column was standardized while irrelevant columns like the quality code measurement including RPE_QC, WSE_QC, WLM_GSE_QC, RPE_WSE_QC and GSE_WSE_QC, were removed, to focus on predictive GWL features. The station column with incomplete or null values were filtered out, as the remaining station records were groups into A to E, enabling detailed analysis across specific locations. Additionally, transformation using standardization technique was done on the selected and prepared data ensuring a consistent scale of variables for accurate model training. The pre-processed and transformed data are further explored in the next subsection.

3.7 Descriptive Data Analysis for Stations

At this stage, the primary focus was to analyse the prepared data, understand relationships between the features, identify trends, and uncover patterns, essential to GWL forecasting. In the preprocessing stage, the data was divided into five stations group (A to E). Each station group retained more than 10,368 records and 6 columns including the MSMT_DATE, WLM_RPE, WLM_GSE, RPE_WSE, GSE_WSE and WSE. Various comparison analysis was conducted, for instance, Figure.4 analysed the surface and depth of Water of multiple stations over time. This helped understand the difference trends and seasonal fluctuation over the years. Figure 5 also compared the variations using a box plot of ground surface water elevation to the depth of water below ground surface in all Stations. This showed each station geographical and hydrological characteristic. Correlation analysis using heatmap in Figure 6 revealed the interrelationship between features at each station, which can be useful for predictions. Also, Figure. 7 applied time series analysis using seasonal decomposition of trends, seasonality, and residual components. This made way for a full view into the fluctuations and cyclic patterns that influenced the GWL on multiple stations. This analysis showed valuable insights into the hydrological behaviour of different stations, providing robust knowledge for building predictive time series models. The subsequent sections use these insights to build the models proposed for this research.

Figure 4: Comparison Analysis Based on Surface and Depth of Water on Multiple Stations over Time

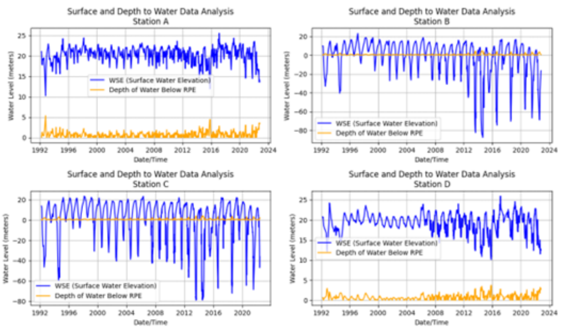


Figure 5: Comparison Analysis based on Ground Surface Elevation and Depth Below Ground Surface in all stations.

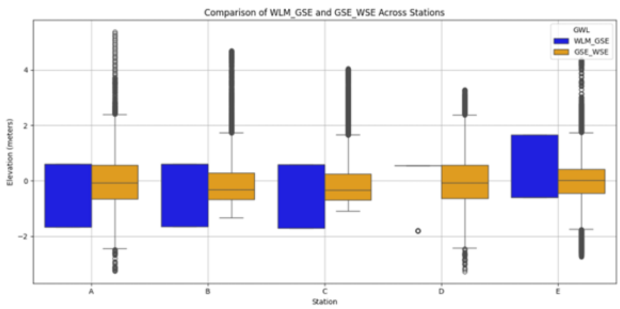


Figure 6: Correlation Analysis of Multiple Stations

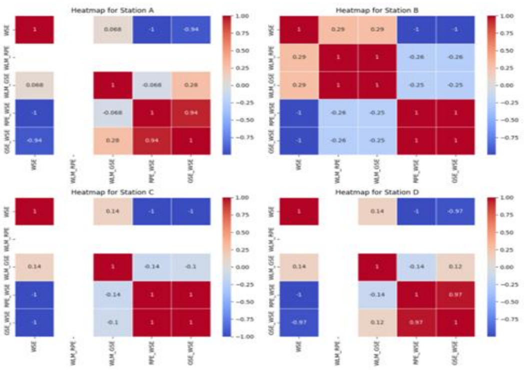
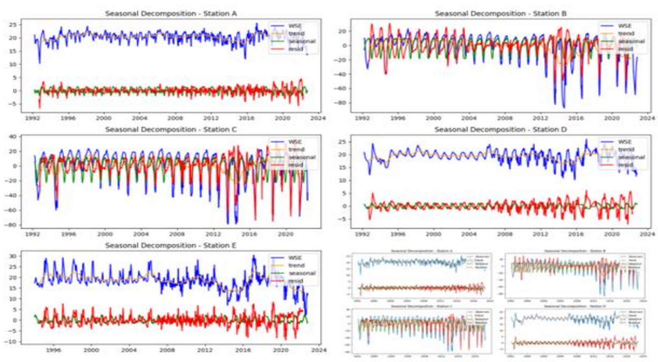


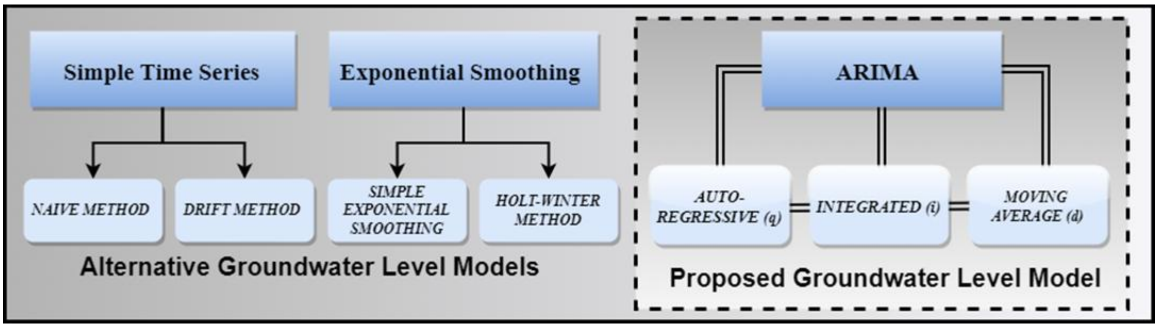
Figure 7: Seasonal Decomposition of all Stations over Time.



3.8 Time-Series Modelling Approach for Groundwater Levels Forecasting

This research uses time-series models to forecast the historical GWL data as illustrated in Figure 8. The approach is split into two. The first part explores four alternative models: Simple Time Series Models like Naïve and Drift methods which make straightforward forecasts from past values and Exponential Smoothing Models like simple Exponential Smoothing and Holt-Winter methods which capture short-term forecasts of both trends and seasonal variations. The second part focuses on the proposes ARIMA model (Autoregressive Integrated Moving Average), with its ability to manage more difficult and temporal structure problems. ARIMA consists of Auto-regressive (AR or q) which uses dependencies from both past and current values, Integrated (I) for trends differencing, and Moving Average (MA or d) for smoothing noise based on past error values. The ARIMA model is a more flexible choice and from previous literature, it has proven to capture trends and seasonal patterns (Monir et al., 2023). Regardless of this, the ARIMA model would be compared with the alternative models with the aim of finding a best-fit model that can answer the question in section 1.1. The method by which the best-fit model is evaluated is explained in the following subsection.

Figure 8: Time-Series Models for Groundwater Level Forecasting



3.9 Evaluation Metrics for Groundwater Levels Forecasting

The evaluation metrics used to analyse the alternative and the proposed models in Figure 8 are explained in this section. These are sets of statistical metrics that measure the accuracy and reliability of the GWL models. They help in quantifying the model's performance by comparing them against a set of actual observed values from the data. The evaluation metrics include coefficient of determination (R^2), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Akaike's Information Criterion (AIC), and Mean Absolute Percentage Error (MAPE). Also, based on the objectives in section 1.2, each model would further be compared using RMSE and R^2, to find out the most efficient model based on their accuracy and predictive power. The characteristics of each evaluation metric are further explained in Table 3, which are applicable to groundwater level forecasting (Iqbal et al., 2021).

Table 3: Evaluation Metrics for Time Series Models in Groundwater Levels Forecasting

Metrics	Characteristics	Interpretation	Formula
Coefficient of Determination	Goodness-of-fit – Measures how well the model fits the data	Closer to **1** or **> 0.7** indicates better model fit	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$
Root Mean Squared Error	Model accuracy – Measures the average magnitude of error, sensitive to large errors	Lower RMSE of **< 1.5** indicates more accurate predictions	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
Mean Absolute Error	Model accuracy – Measures the average absolute difference between actual and predicted values	Lower MAE of **< 10%** indicates higher prediction accuracy	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
Mean Absolute Percentage Error	Model Accuracy – Measures the average percentage error between actual and predicted value	Lower MAPE **< 10%** **(**Excellent) or <**50%** (Acceptable) indicates higher forecast accuracy	$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i - \hat{y}_i}{y_i}\right\rvert \times 100$
Akaike's Information Criterion	Model complexity – Measures the relative quality of statistical models for a given set of data	Lower AIC values suggest a simpler, better fitting model	AIC = 2k-2 In(L)

3.10 Conclusion of Methodology

The section provides a structured multifaced framework in GWL forecasting. It has outlined each process applied from data collection, preprocessing, transformation, analysis, modelling and evaluation metrics used. The ARIMA model is adopted as the primary model and compared with the Exponential Smoothing and Simple Time Series models, to determine the best-fit model for sustainable water resource management. The methodology framework guarantees the integrity of the data as it systematically addresses data quality issues, incorporating exploratory and data analysis and applying different standardization techniques to further improve the model performance to support the research question.

4 Implementation and Experimental Setup

This section presents the experimental setup employed to implement the proposed methodology, which encompasses several multi-layered processes involved in building, training and testing the models in forecasting groundwater levels. With the use of real-world applications and tools as depicted in Figure. 2. key resources, feature selection process, and data preparation (Validation and splitting), for time series modelling are further explained. This effort ensures that the implementation and experiments done provides a comprehensive framework aligned with the research objective in section 1.2.

4.1 Groundwater Level Resources and Tools

The experiments were performed on Jupyter Notebook within the Anaconda package management as an integrated development environment. The package uses Python as the primary Programming Language to conduct extensive coding for data manipulation, processing and analysis. The environment installed and updated necessary core Python libraries like Pandas, Seaborn, NumPy, Matplotlib for data handling and visualization while Sklearn and Statsmodels were employed for modelling. Furthermore, MongoDB worked as a database software to store and manage the large GWL dataset, while Microsoft Excel created a project timeline to track each deliverable set as objectives. Draw.io was used as a diagrammatic tool to create visuals for the research methodology, and workflow process. The Intel (R) Core i7- -2640M CPU @ 2.80GHz and 8.00 GB RAM were components of the operating system used to run the experiments.

4.2 Feature Selection of Stations

Each station group (A-E) retains six core features including MSMT_DATE, WLM_RPE, WLM_GSE, RPE_WSE, GSE_WSE, and WSE. Recognizing feature selection as an important step in any time series modelling, a simple yet effective approach was used, as non-essential features (Column) were dropped in other to provide significant information for GWL forecasting (Iqbal et al., 2021). Understand that the main goal of the research is to conduct a time series analysis that can forecast WSE above sea level over an eight-year period, hence only the MSMT_DATE and WSE features were finally retained. By conducting several correlation analyses shown in Figure 6 and seasonal decomposition shown in Figure. 7 on all station groups, the selected features can be justified based on the following reasons.

Target Variable Focus – WSE column represents a key variable for forecasting the water surface above ground level, hence retaining it ensures the model focuses directly on the prediction of the most relevant outcome when it comes to groundwater management.

Relevance of Time Dimension – MSMT_DATE offers a temporal structure that time series can rely on as it captures historical trends, seasonal patterns, and long-term fluctuations important for prediction.

Irrelevant Exclusion – Features such as WLM_RPE, WLM_GSE, RPE_WSE, GSE_WSE, added contextual hydrological information but had minimal predictive value unlike WSE, hence their exclusion

Dimensionality and Efficiency – Reducing the dataset to WSE and MSMT_DATE, simplified the analysis, by gaining computational efficiency and avoided overfitting from redundant data

4.3 Data Validation and Splitting

This stage is the last phase before developing the models. The validity of the data was explored by checking the class distribution of each station, A through E, to ensure data balancing. The MSMT_DATE and WSE column was confirmed to be in a proper datetime and float64 format. Furthermore, the data of each station (A-E) was split into train and test sets using a ratio of 80:20 per cent respectively. This approach provided a very robust foundation for training models on past patterns while reserving the unseen data for model evaluation

4.4 Conclusion of Implementation and Experimental Setup

The implementation and experimental setup have efficiently established a base framework for building time series models for groundwater levels. With the use of robust tools, features selection and validation and splitting of the data to ensure integrity, this research was set to achieve its aim of obtaining an accurate forecast of water surface elevation. This approach enhances the reliability of the subsequent modelling efforts, while also being in line with the objectives.

5 Evaluation Result and Analysis

This section evaluates and interprets the results of forecasting GWL using various time series models. The analysis aimed to assess the performance based on multiple statistical metrics shown in Table 3, while investigating their implications. In the process, the research tests three different experimental cases – Simple Time Series, Exponential Smoothing, and ARIMA models, while analysing their performance and contributions to the research question in 1.1.

5.1 Experiment 1: – Simple Time Series Models

Table 4: Performance Evaluation Results Using Naïve Base and Drift Methods

STATIONS	Naive Base Method					Drift Method				
	RMSE	MAE	R^2	MAPE	AIC	RMSE	MAE	R^2	MAPE	AIC
A	2.39	1.76	-0.11	9.51	3840.23	2.40	1.77	-0.12	9.56	3865.76
B	56.98	53.55	-6.30	3449.07	17555.15	65.07	62.00	-8.52	3772.80	18132.04
C	22.87	14.35	-0.30	334.59	13584.85	22.69	14.24	-0.28	326.94	13551.08
D	5.87	5.11	-2.20	25.16	7487.49	6.58	5.85	-3.01	29.02	7968.60
E	5.34	3.83	-0.00	35.43	6952.80	5.39	3.87	-0.02	35.97	6989.77

The first experiment used simple time series models including naïve base and drift methods for its analysis. The results shown in Table 4 are based on Table 3, across station A to E. From the performance of the naïve base method, it generally had lower error metrics (RMSE and MAE), for most stations compared to the drift method. Also, the R^2 values across all stations are negative, indicating poor fit and the inability to explain the variance of the data. Relatively, moderate values of MAPE are obtained for stations A, D and E, while B and C station values obtained are very high, as these show significant inaccuracies in forecasting. The drift method has higher RSME and MAE across all stations, which means that it presents greater forecast errors compared to Navie method. Negative R^2 values are more evident in this method, specifically for station B with -8.52 and station D with -3.01, which implies a poorer fit of the model compared to the naïve method. The MAPE values across all stations are very high, especially for stations B and C with 3772.80% and 326.94% respectively, which also indicate a high inaccuracy while forecasting. Overall, the naïve base method outperformed the drift method indicating lower RMSE, MAE, and AIC values for most stations, while both models struggle with poor R^2 and high MAPE, especially in the case of stations B and C. Stations A, D, and E showed relatively better results. The naive base method may have had slightly more reliable forecasts in predicting groundwater levels, but both models serving as baseline references were not sufficient in forecasting GWL due to their inability to capture temporal patterns.

5.2 Experiment 2: – Using Exponential Smoothing Models

After analysing the simple time series methods and their limitations, experiment 2 explored more sophisticated models including simple exponential smoothing (SES) and Holt-winter methods to improve the forecasting accuracy. Unfortunately, the results derived using these two methods indicated significant challenges in providing reliable forecasts, especially in the eight-year prediction. SES performance metrics like RMSE and MAE were considerably high in most stations (for example, RMSE for station B was 56.98 and for station C was 22.87). The R^2 values were negative for all stations, as this indicates a poor fit model, while MAPE was undefined, because of its irregularities in handling the GWL data or extreme forecast error. As shown in Table 5, the forecasted value was static and failed to capture variability, indicating that SES was not able to adapt to the inheritance fluctuations of groundwater levels. However, Holt- Winter method extended SES by considering both trends and seasonality components, but did not perform even as well as SES. For example, during the analysis of Holt-Winter, it RMSE values in station B blew to up 553.99 and 126.74 for station A indicating extreme deviations from the observed data. Also, it had negative R^2 values such as -3116.18 for station A, showed that the method did worse than even the simple Naïve base method. Table 5 also reflects how the future forecast of the eight years period further moves away into unrealistic values which does not consider or support hydrological realities. Overall neither SES nor Holt-Winter methods significantly improved the performance of the basic time series models. Instead, their inability to cope with the complexity of GWL data suggests these models are unsuitable for long-term forecasting of groundwater levels.

Table 5: Performance Forecast of Simple Exponential Smoothing and Holt-Winter Methods

STATIONS	Exponential Smoothing Forecast Years				Holt-Winter Forecast Years			
	1 5	2 6	3 7	4 8	1 5	2 6	3 7	4 8
A	21.325	21.325	21.325	21.325	11.313	-25.218	-61.823	-98.355
	21.325	21.325	21.325	21.325	-134.887	-171.419	-208.014	-244.546
B	-59.627	-59.627	-59.627	-59.627	-91.347	-237.503	-384.213	-530.368
	-59.627	-59.627	-59.627	-59.627	-676.524	-822.680	-969.396	-1115.551
C	11.520	11.520	11.520	11.520	6.850	-16.409	-39.912	-63.172
	11.520	11.520	11.520	11.520	-86.432	-109.692	-133.170	-156.430
D	14.030	14.030	14.030	14.030	11.317	8.501	5.668	2.852
	14.030	14.030	14.030	14.030	0.036	-2.780	-5.620	-8.436
E	18.174	18.174	18.174	18.174	34.280	82.037	130.120	177.877
	18.174	18.174	18.174	18.174	225.634	273.391	321.450	369.207

5.3 Experiment 3: – Proposed ARIMA Model

This experiment further used the proposed ARIMA model. It was expected to capture and improve temporal dependencies for the complex GWL data. The model showed improvements over the simple time series and exponential smoothing models in Table 6 but still faced difficulties in the accurate forecasting of groundwater levels. For instance, although the RSME and MAE were relatively lower in some stations than in previous methods, significant errors were still recorded, especially in stations B and C with RMSE of 65.79 and 23.19 respectively. Also, this model consistently followed the baseline models with negative R^2 values across all stations indicating a poor performance in explaining the variance in groundwater data. Based on the forecast patterns, table 6 shows that most stations' predictions were either stable or showed a simplistic trend, like a gradual increase in Station C or a decline in Station D, which also failed to capture the complex dynamics. Long-term forecasts for some stations, such as Station B, were erratic, showing instability in model predictions.

Table 6: Performance Evaluation and Forecast Results for ARIMA Model

STATIONS	ARIMA MODEL								
	Evaluation Values					Forecast Value (Years)			
	RMSE	MAE	R^2	MAPE	AIC	1 5	2 6	3 7	4 8
A	2.31	1.74	-0.04	NAN	26574.8	21.026 21.026	21.026 21.026	21.026 21.026	21.026 21.026
B	65.79	62.39	-8.73	NAN	17385.5	-68.972 -69.018	-69.018 -69.018	-69.018 -69.018	-69.018 -69.018
C	23.19	14.56	-0.34	NAN	22500.2	12.162 12.171	12.171 12.171	12.171 12.171	12.171 12.171
D	6.05	5.28	-2.39	12.61	-23922.9	13.816 13.816	13.816 13.816	13.816 13.816	13.816 13.816
E	5.34	3.83	-0.04	-0.00	-256.01	18.260 18.260	18.260 18.260	18.260 18.260	18.260 18.260

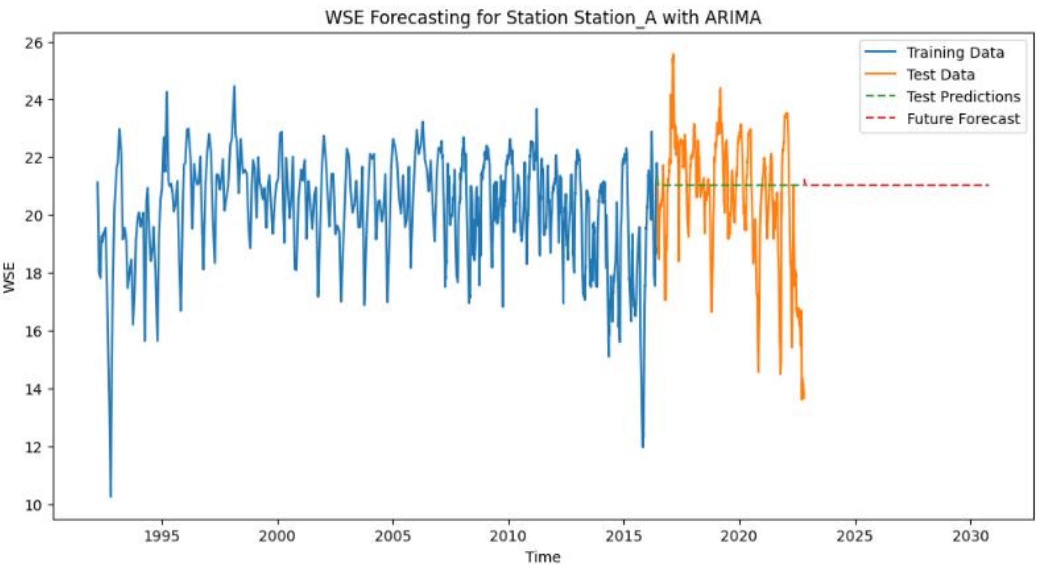


Figure 9: WSE Eight Years Forecast for Station A using ARIMA Model.

In figure 9, the ARIMA model trains the data in blue, as it captures the historical values of WSE. Fluctuations are observable in this trend, which are gradually stabilized over time. However, the test data represented by the orange line shows a series of high variability is not well predicted by the model. The test predictions in the green line remain almost constant throughout, reflecting none of the fluctuations in the test data. Finally, the eight years future forecast represented in red dashed line is completely flat (straight) and constant through 2030. This shows that the ARIMA model failed to predict and forecast the actual groundwater data.

5.4 Comparisons Analysis of Models

As part of the objective to compare all models, the three experiments focusing on the statistical performance of RMSE and R^2 metrics exhibited various degrees of their predictive powers. The ARIMA model turned out to be the best-performing model. For instance, across all stations and methods, station A achieved the lowest RMSE value of 2.31 and the highest but yet negative R^2 of -0.04. The simple time series methods showed basic limitations in their predictive capabilities while exponential smoothing provided a middle ground in capturing smooth trends as it struggled with variability. Despite the relatively better performance of ARIMA results, it was well below the accuracy level that was anticipated for groundwater research. The ARIMA model based on literature, proposed a robust tool that is flexible, interpretable and suitable to capture long term trends for sustainable groundwater management (Takafuji et al., 2019). Overall, the ARIMA model showed some promises, but still, none of the models gave satisfactory accuracy, hence these approaches are limited for complex groundwater forecasting.

6 Discussion

This study investigated the potential of time series models in GWL forecasting for sustainable water resource management. During the analysis, the research tried to fill in the gaps from previous related works (in subsection 2.3) as it proposed the ARIMA model in handling long-term trends. However, the model fell short of expectations. This underlines the complexity of the groundwater system, possibly requiring the adoption of a more advanced approach. A number of challenges were faced during the investigations, as this may have influenced the suboptimal results of the models. In the data collection phase (in subsection 3.4), it was difficult to know the ideal data resolution (daily vs. monthly) suitable for forecasting. The data preparation phase (in subsection 3.6) also had issues in terms of 'stations' grouping and categorisation, as only five stations within 600 locations were selected. A more focused approach to grouping might have been more insightful. Furthermore, not many studies conducted groundwater predictions using simple time series or exponential smoothing, meaning that there were limited opportunities for benchmarking and putting the results into perspective.

Despite all its challenges, this research followed a structured methodology pipeline (in subsection 3.1) that included robust pre-processing and transformation steps of GWL data. This pipeline ensured data quality and a minimum amount of noise. However, the inherent limitations of the models were reflected by their inability to adapt to dynamic groundwater fluctuations. Therefore, it seems that the hypothesis of time series models being able to forecast GWL accurately, for sustainable water resource management is partially unsupported. The ARIMA model performed relatively better; however, none of the models were found to be good enough to use with confidence in decision-making. For instance, a real-world interpretation of the best results previously presented in Figure 9, implies that while the ARIMA model could provide a basic estimate of average groundwater levels, it cannot be used to predict the increase or decrease of water levels, which are important for resource planning. The flat forecast from the image further suggests that the model does not consider seasonal fluctuations, droughts, or human-made influences, and therefore, this approach is less useful for dynamic strategies in groundwater management.

7 Conclusion

In conclusion, this research has highlighted the limitations of the traditional time series models for GWL forecasting, as it emphasized the need for an advanced approach. While the results of the analysis cannot provide strong evidence to support the research question (in section 1.1), they do offer insights into the limitations applied and avenues for further improvements. This paper contributes to the knowledge base, by documenting the challenges and outcomes of the study, which can help future researchers avoid similar pitfalls and encourage the exploration of more innovative methods. Future work should explore hybrid models while combining ARIMA trend analysis strengths with machine learning to capture complex patterns. In addition, external variables such as climate data and extended datasets could improve model accuracy and predictions. This approach retains ARIMA interpretability while addressing its limitations.

Acknowledgement

I am thankful to my supervisor Dr David Hamill for guiding me through a 12-weeks period of my project at the National College of Ireland.

References

Abdellatif Rafik, Yassine Ait Brahim, Abdelhakim Amazirh, Ouarani, M., Bouchra Bargam, Hamza Ouatiki, Yassine Bouslihim, Lhoussaine Bouchaou and Abdelghani Chehbouni (2023). Groundwater level forecasting in a data-scarce region through remote sensing data downscaling, hydrological modeling, and machine learning: *A case study from Morocco. Journal of Hydrology Regional Studies*, 50, pp.101569–101569.
https://doi.org/10.1016/j.ejrh.2023.101569

Aranguren-Díaz, Y., Galán-Freyle, N.J., Guerra, A., Manares-Romero, A., Pacheco-Londoño, L.C., Romero-Coronado, A., Vidal-Figueroa, N. and Machado-Sierra, E. (2024). Aquifers and Groundwater: *Challenges and Opportunities in Water Resource Management in Colombia. Wate*r, [online] 16(5), p.685. Available at: https://doi.org/10.3390/w16050685

Baydaroğlu, Ö., Yeşilköy, S., Sermet, Y. and Demir, I., 2023. A comprehensive review of ontologies in the hydrology towards guiding next generation artificial intelligence applications. *Journal of Environmental Informatics*, 42(2), pp.90-107. Available at: https://doi.org/10.1016/j.jhydrol.2022.128970

Bedi, S., Samal, A., Ray, C. and Snow, D., 2020. Comparative evaluation of machine learning models for groundwater quality assessment. *Environmental Monitoring and Assessment*, 192, pp.1-23. Available at: https://doi.org/10.1007/s10661-020-08695-3

Boo, K.B.W., El-Shafie, A., Othman, F., Khan, M.M.H., Birima, A.H. and Ahmed, A.N., 2024. Groundwater level forecasting with machine learning models: *A review. Water Research*, p.121249. Available at: https://doi.org/10.1016/j.watres.2024.121249

California Department of Water Resources. (2024). 5 Reasons Groundwater is Vital to Californians. [online]
Available :https://water.ca.gov/News/Blog/2019/March-19/5-Reasons-Groundwater-is-Vital-to Californians#:~:text=Groundwater%20provides%2030%20to%2060

California Water Action Plan. (n.d.).2024. *Actions for Reliability, Restoration and Resilience*. Available at: https://www.waterboards.ca.gov/press_room/press_releases/2014/cwap.pdf

California State Water Resources Control (n.d.).2023 Groundwater Basics. [online] Available at: https://www.waterboards.ca.gov/water_issues/programs/groundwater/gw_basics.html

Chen, H.Y., Vojinovic, Z., Lo, W. and Lee, J.W., 2023. Groundwater Level Prediction with Deep Learning Methods. *Water*, 15(17), p.3118. https://doi.org/10.3390/w15173118

Chen, Z.-Y., Zhang, T., Zhang, R., Zhu, Z., Yang, J., Chen, P., Ou, C.-Q. and Guo, Y. (2019). Extreme gradient boosting model to estimate PM2.5 *concentrations with missing-filled satellite data in China*. 202, pp.180–189.
https://doi.org/10.1016/j.atmosenv.2019.01.027

Ferreira, L.B., da Cunha, F.F., de Oliveira, R.A. and Fernandes Filho, E.I. (2019). Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM: *A new approach. Journal of Hydrology*, 572, pp.556–570.
https://doi.org/10.1016/j.jhydrol.2019.03.028

Gaur, S., Johannet, A., Graillot, D. and Omar, P.J., 2021. Modeling of groundwater level using artificial neural network algorithm and WA-SVR model. In Groundwater resources development and planning in the semi-arid region pp. 129-150. Cham: *Springer International Publishing, Cham*. https://doi.org/10.1007/978-3-030-68124-1_7

Goodarzi, M. (2020). Application and performance evaluation of time series, neural networks and HARTT models in predicting groundwater level changes, Najafabad Plain, Iran. *Sustainable Water Resources Management*, 6(4).
Available at: https://doi.org/10.1007/s40899-020-00427-2.

Groundwater Foundation (2022). What is Groundwater? [online].
Available: https://groundwater.org/what-is-groundwater

Iqbal, N., Khan, A.N., Rizwan, A., Ahmad, R., Kim, B.W., Kim, K. and Kim, D.H., 2021. Groundwater level prediction model using correlation and difference mechanisms based on boreholes data for sustainable hydraulic resource management. *IEEE Access*, 9, pp.96092-96113. Available at: https://doi.org/10.1109/ACCESS.2021.3094735

Khan, J., Lee, E., Balobaid, A.S. and Kim, K., 2023. A comprehensive review of conventional, machine leaning, and deep learning models for groundwater level (GWL) forecasting. *Applied Sciences*, *13*(4), p.2743.
Available at: https://doi.org/10.3390/app13042743

Kontopoulou, V.I., Panagopoulos, A.D., Kakkos, I. and Matsopoulos, G.K. (2023). A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. *Future Internet, [online]* 15(8), p.255. https://doi.org/10.3390/fi15080255

Lall, U., Josset, L., Russo, T., 2020. A snapshot of the world's groundwater challenges, in: Gadgil, A., omich, T.P. (Eds.), Annual Review of Environment and Resources, 45. *Columbia Univ, Columbia Water Ctr, New York, NY 10027 USA*, pp. 171–194. Available at: https://doi.org/10.1146/annurev-environ-102017-025800

José Luis Uc-Castillo, Ana Elizabeth Marín-Celestino, Diego Armando Martínez-Cruz, José Tuxpan-Vargas and José Alfredo Ramos-Leal (2023). A systematic review and meta-analysis of groundwater level forecasting with machine learning techniques: *Current status and future directions. Environmental Modelling & Software*, 168, pp.105788–105788.
https://doi.org/10.1016/j.envsoft.2023.105788

Mbouopda, M.F., Guyet, T., Labroche, N. and Henriot, A. (2022). Experimental study of time series forecasting methods for groundwater level prediction. *[online] arXiv.org*. Available at: https://doi.org/10.48550/arXiv.2209.13927

Md Abdullah-Al Mamun, Abu, Mst Nazneen Aktar, Md Nashir Uddin, Md Saiful Islam, Subodh Chandra Pal, Islam, A., Bari, Idris, A.M. and Venkatramanan Senapathi (2024). Predicting groundwater phosphate levels in coastal multi-aquifers: *A geostatistical and data-driven approach. The Science of The Total Environment*, pp.176024–176024. https://doi.org/10.1016/j.scitotenv.2024.176024

Mirsanjari, M.M. and Mohammadyari, F., 2019, March. Application of time-series model to predict groundwater quality parameters for agriculture:(plain Mehran case study). *In IOP Conference Series: Earth and Environmental Science* (Vol. 127, No. 1, p. 012012). Available at: https://doi.org/10.1088/1755-1315/127/1/012012.

Monir, M.M., Sarker, S.C. and Islam, M.N., 2024. Assessing the changing trends of groundwater level with spatiotemporal scale at the northern part of Bangladesh integrating the MAKESENS and ARIMA models. *Modeling Earth Systems and Environment*, 10(1), pp.443-464. Available at: https://doi.org/10.1007/s40808-023-01794-3

Paul, M., Kumar, A., Kumar, H., Viji Amutha Mary, S. Jancy and Grace, J. (2024). Groundwater Quality Prediction using Deep Learning. pp.917–923.
https://doi.org/10.1109/icoeca62351.2024.00162

Rajaee, T., Ebrahimi, H., Nourani, V., 2019. A review of the artificial intelligence methods in groundwater level modeling. J. *Hydrol*. 572, 336–351.
https://doi.org/10.1016/j.jhydrol.2018.12.037

Rashid, A., Alamgir, M., Ahmed, M.T., Salam, R., Islam, A.R.Md.T. and Islam, A. (2022). Assessing and forecasting of groundwater level fluctuation in Joypurhat district, northwest Bangladesh, using wavelet analysis and ARIMA modeling. *Theoretical and Applied Climatology*, 150(1-2), pp.327–345.
Available at: https://doi.org/10.1007/s00704-022-04160-y

Rahman, A.T.M.S., Hosono, T., Quilty, J.M., Das, J. and Basak, A. (2020). Multiscale groundwater level forecasting: Coupling new machine learning approaches with wavelet transforms. *Advances in Water Resources*, 141, p.103595.
https://doi.org/10.1016/j.advwatres.2020.103595

Seyler, H., Gibson, K., Kanyama, Y. and Witthüser, K., 2020. Machine learning models for groundwater availability. *Water Research Commission, Pretoria. WRC Report No. TT*, 845, p.20. https://www.wrc.org.za/wp-content/uploads/mdocs/TT%20845%20final%20web.pdf

Shaikh, M., Birajdar, F., Shaikh, M. and Birajdar, F. (2024). Artificial intelligence in groundwater management: Innovations, challenges, and future prospects. *International Journal of Science and Research Archive, [online]* 11(1), pp.502–512.
https://doi.org/10.30574/ijsra.2024.11.1.0105

Sarma, R. and Singh, S.K. (2022). A Comparative Study of Data-driven Models for Groundwater Level Forecasting. *Water Resources Management*, 36(8), pp.2741–2756.
https://doi.org/10.1007/s11269-022-03173-6

Takafuji, E.H. de M., Rocha, M.M. da and Manzione, R.L. (2019). Groundwater Level Prediction/Forecasting and Assessment of Uncertainty Using SGS and ARIMA Models: A Case Study in the Bauru Aquifer System (Brazil). *Natural Resources Research*, 28(2), pp.487–503. https://doi.org/10.1007/s11053-018-9403-6.

Tyralis, H., Papacharalampous, G. and Langousis, A. (2020). Super ensemble learning for daily streamflow forecasting: large-scale demonstration and comparison with multiple machine learning algorithms. *Neural Computing and Applications*.
https://doi.org/10.1007/s00521-020-05172-3

Wetzel, M., Kunz, S., Maximilian Nölscher, Schulz, A., Biessmann, F. and Broda, S. (2024). Advancing water resource management: *Insights and implications from global machine learning models in groundwater prediction*. https://doi.org/10.5194/egusphere-egu24-11647.