JOURNAL OF ENGINEERING RESEARCH AND TECHNOLOGY, VOLUME 12, ISSUE 2, ACTOBER 2025

Forecasting of Groundwater Total Dissolved Solids in Khan Younis City

Asem H. Shurrab[1], Abdelkareem M. Alashqar[2] and Ashraf Y. Maghari[3]
[1,2,3]Faculty of Information Technology, Islamic University of Gaza, P.O. Box 108, Gaza, Palestine
[1]tasemsh@gmail.com, [2]aashgar@yahoo.com, [3]amaghari@iugaza.edu.ps
[2]Corrsponding author

Abstract— Wells of groundwater is the main resource of water supplies in Gaza Strip. The elevated water consumption, the rainfall shortage and the intrusion of seawater in some aquifers increases the creation of total dissolved solids (TDS) in water. The early forecasting of TDS levels in wells provides more timely information for better management of groundwater resources. In this paper we applied various techniques on timeseries data to forecast the TDS levels of wells in Khan Younis city as a case study. For this purpose, annual timeseries data were collected for seven wells from main two sources to predict TDS amounts for five years in advance. Five different forecasting models which are autoregressive integrated moving average (ARIMA), k-nearest neighbors (KNN), linear regression (LR), random forests (RF) and artificial neural networks (ANN) were applied individually on each well data. According to the mean absolute percentage error (MAPE) metric, the best results of the applied models on six wells (Ahrash, Ayia, El Amal, Riada City, UN Khan Younis and Islamic Relief) is ARIMA (MAPE ratios is between 3% and 9.8%). While the RF model has the best MAPE ratio (5.5%) in the seventh well (Abu Khaled). The results demonstrated that in the next five-year period of each well according to its horizon, the TDS ratio will increase for Ayia, Riadia City, UN khan Younis, Ahrash and Abu Khalid in comparison with the previous five-year period by 22.1%, 5.6%, 4.4%, 2.7% and 1.6% respectively. In contrast, the TDS levels in Islamic Relief and El Amal wells will decrease by 7.2% and 2.1% respectively. The results of this study contribute in better management, planning and risk minimization of water quality.

Index Terms— Time series forecasting, total dissolved solids (TDS), ARIMA, k-nearest neighbors (KNN), linear regression (LR), random forests (RF), artificial neural networks (ANN), MAPE.

I INTRODUCTION

Providing quality water is very important for reserving human lives. The total dissolved solids (TDS) is a type of water pollution that represents the measurement of organic and inorganic molecular materials in water. The nitrates, carbonates, chloride, phosphates, calcium, sodium, potassium and magnesium are mostly the constituents of TDS in water [1]. TDS is often measured in milligrams per liter (mg/l) where the high level of TDS is an indicator of the high amount of contamination in water supplies [2].

The very limited water resources in Gaza Strip increases the risk of providing healthy drinking water. Gaza Strip depends heavily on groundwater wells in supplying water for citizens. The increased consumption due to the overpopulation, the rainfall shortage and the intrusion of the sea and sewage water mostly contributes in the increasing levels of TDS and salinity in groundwater aquifers [3]. The authors in [4] argued that the large amount of salt and the elevated pollution with nitrates reduces the quality of drinking water that directly affect the human health and life. So, the earlier forecasting level of TDS gives important insights for better management of water resources in Gaza Strip especially groundwater wells.

Some governmental and private institutions in Gaza Strip collect periodical data about a variety of groundwater resources such as registering the amount of water consumptions used for human and agriculture. They often take semi-annual water quality measurement for groundwater characteristics. However, there is still a lack of research studies utilizing these useful collected data in order to provide remarkable results for better water quality management.

Forecasting TDS level in water is a type of time series data analysis and modeling. The main objective of time series modeling is collecting and analyzing the past time observations in order to understand the relationships within data of time series then building a model to forecast (predict) useful future values [5].

While there are many studies have been conducted on drawing forecasts and predictions of TDS amount in groundwater using regression models worldwide, very limited studies have been achieved in Gaza Strip.

However, a variety of models (techniques) can be used for conducting TDS forecasts depending on time series data. These models include, for instance, traditional statistic models, stochastic regression models, machine learning models and deep learning models. While there is no universal consensus on categorizing these models, they can be approximately classified as traditional statistic-based models and machine learning (ML)-based models [6].

In this paper we used five different models which are autoregressive integrated moving average (ARIMA), k-nearest neighbors (KNN), linear regression (LR), random forests (RF) and artificial neural networks (ANN) for predicting TDS amount in groundwater wells located in Khan Younis area. Then, the best performing model will be used to forecast the TDS level for five-years period in advance.

The main contribution of this paper is summarized as follows:

- Build a time series dataset for TDS level of groundwater. This was achieved by collecting and combining data from two main sources for seven wells in Khan Younis City that were carefully preprocessed to be used in the potential experiments.
- Examine and compare the performances of five selected different forecasting models utilizing the built and preprocessed dataset.
- Select the best performing model to draw TDS forecasts for the next five years based on the collected data horizon of each well.

The results of this paper provide useful insights and recommendations for better water quality management and better water demand planning that urgently needed by formal institutions in Gaza Strip.

The rest of this paper is organized as follows: a review of previous related studies is provided in Section II. A multistep designed method for performing the objective of this paper is introduced in Section III. Section IV provides the results of our experiments, evaluations and discussions. Section V concludes the paper and recommends for future research orientation.

II Related Work

There is a considerable number of research works conducted on time series forecasting. These works have been done for different domains utilizing various forecasting models.

Regarding the models used, the literature introduced traditional statistical models such as Theta and Box-Jenkins [7], stochastic models such as ARIMA, neural networks such as artificial neural network (ANN) and time lagged neural networks (TLNN) [5], support vector machines (SVM) [8], machine learning (ML) models such as k-nearest neighbors (KNN) and random forests (RF) [2], deep learning models such as recurrent neural networks (RNN) and long short term memory (LSTM) [9].

With regard to the examined domains, the author in [10] used decision trees bagging and RF to build forecasts for stock price directions in clean energy companies, and the authors in [11] used the models of ARIMA, KNN, logistic regression (LR) and support vector regression (SVR) to forecast the energy consumptions in three different areas, while the authors in [12] drew forecasts for carsharing customer demands using the models of regression, moving averages and neural networks. Additionally, the authors in [13] utilized KNN and SVR for traffic flow prediction and the authors in [14] applied ML for forecasting recovery rates on non-performing loans for retail clients. Moreover, the authors in [15] proposed a methodology that applied clustering and forecasting techniques to predict customer behavior to improve marketing.

Focusing on the research studies related to water, there are considerable research works that focused on forecasting the amount (level) of water as well as its quality using salinity and TDS ratio as indicators.

Regarding water level, the authors in [16] applied the feed forward neural networks (FNN) to predict the level of eight wells in South Korea by checking river level and pumping averages of extracted wells as effecting related input variables. And the researchers in [17] used various structures of FNN to forecast groundwater wells located in Montgomery, USA. They used humidity, precipitation, and temperature as input variables and found that the Levenberg Marquardt (LM) training algorithm outperforms the other structures of FNN algorithms. In addition, the authors in [18] compared between SVR and a model named nonlinear autoregressive with exogenous inputs and artificial neural network (NARX-ANN) in forecasting groundwater level (GWL) for wells used in irrigation located in south east of USA. They conducted their assessment based on a mixture of three inputs: daily GWL, precipitation and evapotranspiration, and found that the precipitation is the best optimal mixture for the model input where the SVR outperforms the ANN.

With regard to water quality, the authors in [19] developed a model based on ANN to predict groundwater salinity that is expressed by TDS using the measurement of pH seasonal corps that used well water for irrigating. They used a data sample of 38 wells located in Alexandria, Egypt where considerable results were obtained. And the authors in [20] applied KNN to predict influent flow rate and water quality. The number of nearest neighbors (NNs) was tested under dry and wet weather conditions using the root mean square error (RMSE). They found that square root-based (SR) technique outperforms the distance factor-based (DF) technique in determining the number of NNs. The influent flow was predicted using standard deviation and the water quality was predicted using mean absolute percentage error (MAPE). They produce considerable results for seven-day forecasting with less than 5% for prediction accuracy diffidence.

Also, the researchers in [21] built a model named fast orthogonal search model (FOS) to forecast groundwater salinity in area of Deltona, Florida by collecting data about 27 wells. The forecasting takes into account the relationship between chloride concentration and water demand. They used root mean squared error (RMSE), standard deviation ratio (RSR), Nash Sutcliffe efficiency coefficient (NSEC) and the correlation coefficient R to measure accuracy where FOS gave considerable results for forecasting groundwater salinity.

Additionally, the authors in [22] applied machine learning models to determine the groundwater quality by forecasting their physical components including TDS, potential salinity (PS), sodium adsorption ratio (SAR), exchangeable sodium percentage (ESP), magnesium adsorption ratio (MAR), and the residual sodium carbonate (RSC) parameters through electrical conductivity (EC), temperature (T), and pH. They evaluated the models of adaptive boosting (Adaboost), RF, ANN, and SVR models using 520 samples of data related to fourteen Groundwater quality parameters in Berrechid aquifer, Morocco. The results showed that Adaboost and RF models gave higher prediction performances than did the SVR and ANN.

Moreover, the authors in [23] conducted a comparative study on models of groundwater salinity in Yongnian China. They constructed a series of fitting and prediction models of groundwater salinity based on grey theory and MATLAB on samples of data collected for three wells in the period 2010-2015. Their model gave considerable results in demonstrating a high-level increase of salinity in the examined wells.

Also, the researchers in [24] developed a hydrogeological model to forecast the intrusion of saltwater of coastal zone in Vietnam. They indicated that the freshwater areas are intruded with proportion to the exploiting reserve. While the authors in [25] applied genetic programming (GP) to determine the relations between groundwater quality parameters such as total hardness (TH), TDS and electrical conductivity (EC). They compared the GP with ANN and adaptive neuro-fuzzy inference system (ANFIS) and demonstrated that GP outperforms other methods in generating compact models that are structurally independent. Moreover, they showed that GP is more robust un estimating values of water quality parameters despite the high degree of nonlinearity among various hydro chemical properties.

Related to the area of Gaza Strip, the authors in [26] applied the models of forecasting, linear regression and multiple regression to forecast salinity using the chloride ion-selective as indicator. Their results stated that the increase of salinity will cover 95% of the selected area in Gaza Strip.

A similar study to our one has been conducted by [27] where the authors used multiple forecasting algorithms to predict the salinity rates and water levels of the groundwater in Deir El-Balah area, Gaza Strip. They applied each of exponential smoothing (ETS), state space model with Box–Cox transformation, ARMA errors, trend and seasonal components (TBATS) and ARIMA. They also applied a hybrid model as a combination of ARIMA with each of neural network (NN), ETS, and TBATS models on two datasets collected for some randomly selected wells. One of the datasets was used for forecasting the water levels and the other was used for forecasting the salinity level. Their study demonstrated that ARIMA gave remarkable MAPE scores with comparison to other models of forecasting salinity rates. It also performed better when combining it with NN model. The results confirmed that there is an increase in salinity and a decrease in water level for the examined groundwater wells.

In conclusion, there are many researches handled forecasting of water quality based on salinity and TDS but almost all of these studies conducted on areas other than Gaza Strip. While, and to the best of our knowledge, there is only one study that was conducted in Deir Albalah area for modeling the forecasts of salinity amount of groundwater, our study focusses on forecasting TDS of water in other area of Gaza Strip which is Khan Younis city using different forecasting models. And it is noteworthy that the results of our study add more approved forecasting indicators to help decision makers in reducing risk related to groundwater as a main resource of water in Gaza Strip.

III The Proposed Methodology

This section introduces the proposed designed methodology used in achieving the main purpose of this study. As shown in Figure 1, the method consists of five main processing steps which are: collecting data, preprocessing data, building and examining forecasting models, evaluating and comparing models results and finally creating TDS forecasts. An explanation for these steps is provided in the next subsections.

A Collecting Data

We collected the TDS data from two main sources: The Coastal Municipalities Water Utility (CMWU) and the Ministry of Health (MoH) as reported by their chemical test laboratories. CMWU has TDS data for the period of 1988-2018 that mostly archived in physical papers, while MoH has larger TDS data for the period of 2005-2018 that archived in traditional computer files. It is noteworthy that, in both sources, the characteristics of water extracted from wells including TDS readings are archived two times per a year, one in spring and the other in autumn.

B Preprocessing Data

For increasing the quality of our collected data and to reduce forecasting uncertainty, we applied various important tasks for data preprocessing including integrating data, removing redundancies and conflicts, handling null (missing) values and removing noise.

Data integration includes combining the data that were collected from the two main sources as said before. This may result in some redundancies and conflicts so these potential problems were carefully solved. To prepare more convenient data that will be used for forecasting in our study we applied two main criteria. First, we selected only the wells that have a considerable number (i.e. more than 15) of time series readings over years. Second, we excluded the wells that each has more than 50% of null or missing values in its reading samples. The result of applying these criteria reduces the examined wells from 51 to 7. Table 1 shows the seven selected wells and the number of available readings for each one. The names of the selected wells are: UN Khan Younis, El Amal, Ahrash, Islamic Relief, Riada City, Aiya and Abu Khalid. For handling any missing TDS values, we assigned it by the outcome of calculating the average of the prior and the posterior readings of the well. The number of new readings for each well after handling missing values is shown in the last column of Table 1.

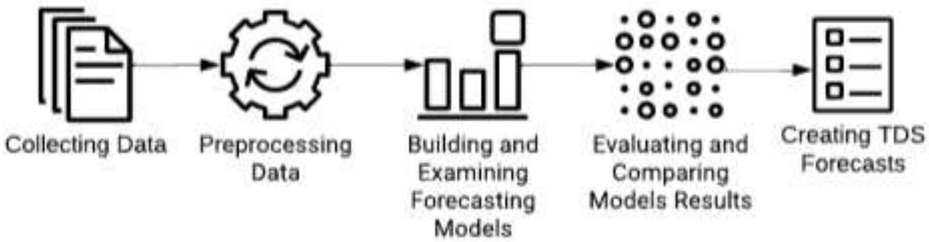

Figure 1: The main steps of the proposed methodology

Table 1: The selected wells in the created dataset

Well Name	Readings Period	No of Readings	% of Null Values	No of Preprocessed Readings
UN Khan Younis	1999-2009	21	4.5%	22
El Amal	1998-2018	39	7.1%	42
Ahrash	1987-2018	55	14%	64
Islamic Relief	2010-2018	18	0%	18
Riada City	2002-2017	29	9.4%	32
Ayia	1989-2018	54	8.5%	59
Abu Khaled	2002-2014	22	12%	25

Since some TDS samples were archived manually, it may include mistakes such as anomalies (outliers) readings. Figure 2 and Figure 3 show instances of anomalies for Riadia City and Abu Khalid wells respectively. Not fixing these mistakes may negatively affect the stability of data dataset and as a result it may undesirably affect the forecasting accuracy. To solve this problem, we manually reviewed the samples for each well and applied the average technique as used above when handling missing values.

Figure 2: Riada City well with anomalies values

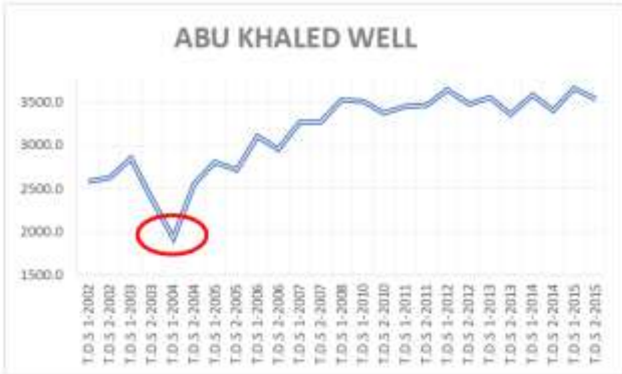

Figure 3: Abu Khalid well with an anomaly value

C Building ad Examining Forecasting Models

We selected five different models to perform TDS forecasting over the built dataset. The selected models are described next.

- **ARIMA**

ARIMA (autoregressive integrated moving average) is one of the most well-known linear time series forecasting models. ARIMA has the ability to detect patterns and relations within time series data. It includes three main elements; autoregression (AR), integration (I) and moving average (MA). The AR element detects the dependencies between an observation point and its lagged values. The I elements makes the values of time series data stationary by removing the seasonality and trends of data. The MA is responsible for the relationship between an observation and the residual errors when using the moving average in the lagged observations. The main limitation of ARIMA is that it assumes the linearity nature of manipulated data [28].

- **K-Nearest Neighbors (KNN)**

The K-nearest neighbors (KNN) is considered a nonparametric regression technique that uses Euclidean distance to measure the similarity between input observations to forecast an output. Specifically, to forecast a value for a point, the KNN finds and picks a number of K (e.g. three or five) other points from the training input observations that are closest neighbors to this point then calculates its prediction as an average of the K neighbors' targets.

- **Linear Regression (LR)**

Linear regression (LR) is a type of linear models that heavily studied and practiced over hundreds of years. LR applies linear function and finds its parameters (e.g. slope and offset) that minimizes the mean square error (difference) between the true input target value and the predicted value within the training points.

- **Random Forests (RF)**

Random forest (RF) is a collection of independent decision trees work in parallel and are created randomly to be slightly distinct. A decision tree is a hierarchy of if/else questions that lead to a predicted decision. The randomization of the trees in RF is achieved by choosing the input observation used in building the decision tree or by choosing the features from the test splits. The final regression prediction of RF is obtained by calculating the average of prediction results produced by each individual tree in the forest [29].

- **Artificial Neural Networks (ANN)**

Artificial neural networks (ANN) – also called multilayer perceptron (MLP), is a collection of thousands of connected nodes (neurons) that organized into three types of layers: input layer, one or more hidden layers and output layer. The nodes of the input layer are connected with the nodes of the hidden layer via links that represent input parameters (weights). The nodes in the hidden layers implement activation functions such as ReLU to transform inputs into predicted output values. The output layer produces a list of probabilities and picks the predicted value that has the lowest difference (error) between true and predicted values. The ANN backpropagates the error value, uses it as feedback and applies optimizer method such as Adam to finetune the weights of the same input observations in order to minimize error. This process is repeated several times (epochs) until reaching the best error value. ANN has the capability of manipulating nonlinear observations [30].

D Evaluating and Comparing Models Results

The performance of the forecasting model is commonly evaluated using the mean absolute percentage error (MAPE) metric that is a widely used for forecasting problems [31]. The MAPE value can be calculated using the following formula:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{T_i - P_i}{T_i}\right| 100\%$$

Where N is the total number of experimented samples by the model algorithm, T_i and P_i are the true (actual) and the predicted values of a sample i respectively. Also, the MAPE value will be used as a comparison metric for the performances results obtained from examined models.

E Creating TDS Forecasts

In this step, when experimenting all of the models using specific well data, the outperforming one will be selected for creating the forecasts of this well for a period of five years in advance.

IV Experimental Results and Forecasts

We experimented each of ARIMA, KNN, LR, RF, and ANN models using our created and preprocessed timeseries TDS data of seven wells located in Khan Younis area. The performance for each model is evaluated using MAPE metric. The dataset is initially split into roughly 80% and 20% for training and testing processes respectively as shown in Table 2.

We used Python programming language for developing the models' algorithms. For KNN, the number of neigbours was set to 10. The ANN is built as sequential layers including one hidden layer that accepts a batch of 128 input with "ReLU" activation function followed by an output layer for predicting one value. The network was compiled using "Adam" optimization function that optimized using mean squared error (MSE) loss function.

We experimented all of the five selected forecasting models on each data of the seven wells. Table 3 shows the MAPE values as produced by each examined model for each individual well data.

Table 2: Dataset splitting into training and testing periods

Well Name	Training Period	Samples	Testing Period	Samples
UN Khan Younis	1999 – 2007	17	2007 - 2009	5
El Amal	1998 – 2014	33	2014 - 2018	9
Ahrash	1987 – 2012	51	2012 – 2018	13
Islamic Relief	2010 – 2016	14	2017 - 2018	4
Riadia City	2002 – 2014	25	2014 - 2017	7
Ayia	1989 – 2012	47	2012 - 2018	12
Abu Khaled	2002 – 2012	20	2012 - 2014	5

Table 3: MAPE performances results using five different forecasting models

Well Name	ARIMA	KNN	LR	RF	ANN
UN Khan Younis	**8.1**	45.3	22	32.2	37.5
El Amal	**7.9**	39.9	15.4	29	19.4
Ahrash	**3**	10.9	6.9	7.1	9.5
Islamic Relief	**11**	39.9	15.5	28.4	16.3
Riadia City	**9.8**	18.2	16.5	14.3	14.7
Ayia	**3.4**	26	9	22.8	22
Abu Khaled	8.9	6.5	12	**5.5**	5.9

As shown in Table 3 the ARIMA model outperformed the other models for six of the seven examined wells according to the MAPE scores. The ARIMA achieved the MAPE values of 3, 3.4, 7.9, 8.1, 9.8 and 11 for Ahrash, Ayia, El Amal, UN Khan Younis, Riada City and Islamic Relief wells respectively. While the RF has the superior value for Abu Khalid well among the other selected models. It is clear that the best MAPE values ranges between 3 and 11 and achieved by ARIMA and RF models. This mainly demonstrates the capability of ARIMA model, and RF in some cases, in forecasting TDS amount using linear time series data.

For each well, the outperformed model was selected to build TDS forecasts for the term of five years in advance for this well. Figure 4, Figure 5, Figure 6, Figure 7, Figure 8, Figure 9 depict the TDS forecasts for UN Khan Younis, El Amal, Ahrash, Islamic Relief, Riada City and Ayia wells respectively using ARIMA model while Figure 10 depicts TDS forecasts for Abu Khalid well using RF model.

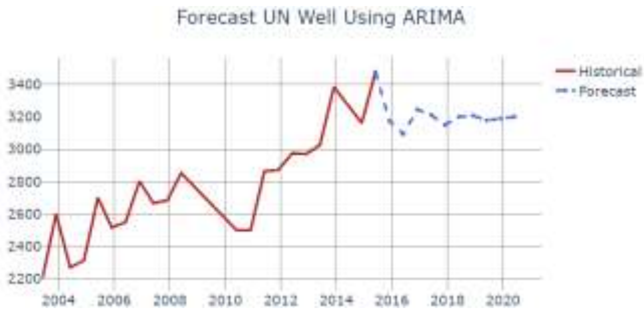

Figure 4: UN Khan Younis well forecast using ARIMA

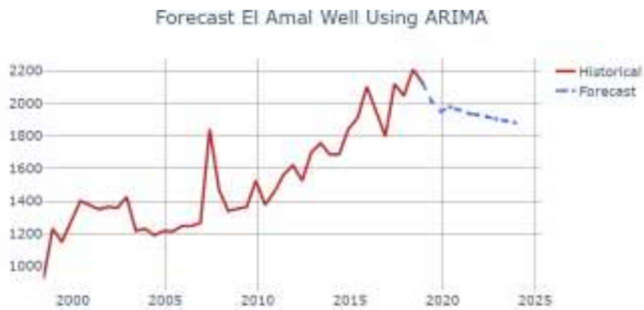

Figure 5: El Amal well forecast using ARIMA

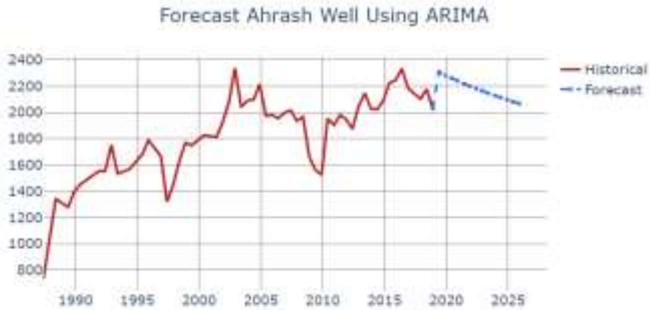

Figure 6: Ahrash well forecast using ARIMA

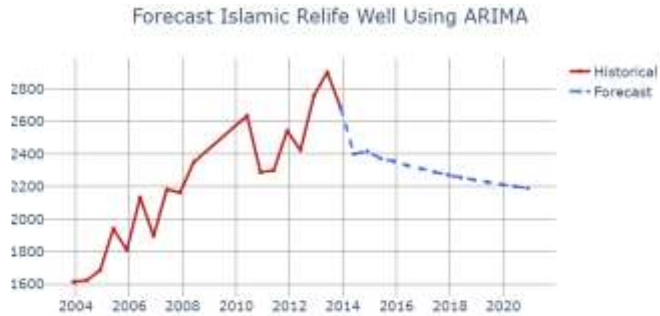

Figure 7: Islamic Relief well forecast using ARIMA

Figure 8: Riadia City well forecast using ARIMA

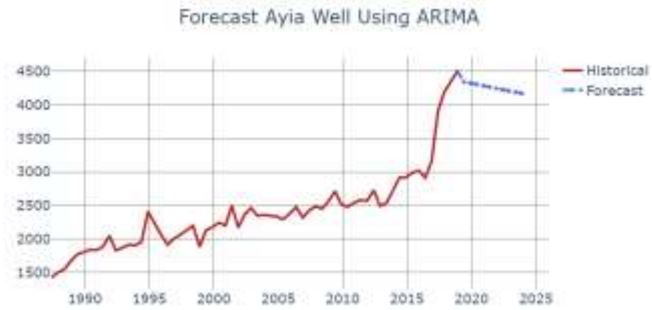

Figure 9: Ayia well forecast using ARIMA

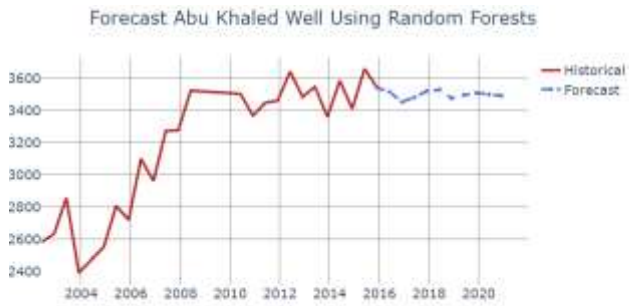

Figure 10: Abu Khaled well forecast using RF

To determine the amount of change in TDS levels, we calculated the difference between the forecasted TDS rates of the next five years and the TDS rates of the past five years.

As shown in Table 4, the forecasted TDS levels of the groundwater are expected to increase in (Ayia, Riadia City, UN khan Younis, Ahrash and Abu Khalid) in comparison with the previous five-year period by 22.1%, 5.6%, 4.4%, 2.7% and 1.6% respectively, compared to the previous five-year period. In contrast, the TDS levels in Islamic Relief and El Amal wells are expected to decrease by 7.2% and 2.1% respectively.

The results in Table 4 indicate that the TDS levels in (Ayia, Riadia City, UN khan Younis, Ahrash and Abu Khalid) are relatively high. This higher TDS rate is due to an imbalance between water consumption and water recharge resulting from the decreasing levels of rainfall over time [32]. On the contrary, the decreases in TDS rates of Islamic Relief and El Amal wells may result from reduced pumping from these two wells and increasing rainfall in the respective areas. It is worth mentioning that Groundwater TDS levels are affected by two primary factors: rainfall, which recharges groundwater, and the pumping mounts from wells. Increased rainfall combined with reduced pumping leads to lower TDS levels [33].

The large values of the forecasted TDS rates in most of the examined wells in this study, give useful indicators and good insights for the decision makers of water distribution in Khan Younis area particularly and in Gaza Strip generally in order to make actions for better management of groundwater supplies.

Table 4: Deviation of TDS rates compared to the last 5 years

Well Name	Previous Period	Average TDS ratio	Forecast Period	Average TDS ratio	Deviation ratio
UN Khan Younis	2011 - 2015	30504.5	2016 - 2020	31852.8	4.4%
El Amal	2014 – 2018	19790.0	2019 – 2023	19356.5	-2.1%
Ahrash	2014 – 2018	2153.7	2019 – 2023	2211.2	2.7%
Islamic Relief	2009 – 2013	25049.0	2014 – 2018	23257.2	-7.2%
Riadia City	2012 – 2016	28772.5	2017 – 2021	30390.0	5.6%
Ayia	2014 – 2018	34875.0	2019 – 2023	42569.1	22.1%
Abu Khaled	2009 – 2013	34605.6	2014 – 2018	35159.7	1.6%

In this study, our experimental results are constrained to TDS timeseries data that were collected periodically twice per a year, and to minimize the effect of relatively high diverge between timeseries readings, it is recommended to archive such samples in shorter time periods (e.g. at most every month) by the responsible intuitions.

V Conclusion and Future Work

In this paper we applied and compared five different models for forecasting the total dissolved solids (TDS) in groundwater wells located in Khan Younis city. The data were collected from two main sources: The Coastal Municipalities Water Utility (CMWU) and the Ministry of Health (MoH) in Gaza Strip then they were preprocessed following several techniques to produce appropriate dataset related to seven wells. The performances of the selected group of models were examined and compared for each well individually using MAPE metric. According to MAPE, the ARIMA model outperformed the other models for 6 of 7 examined wells while the RF achieved the best MAPE value for the seventh well. The results demonstrated that there is an increase of TDS levels in five of the examined wells for five years in advance. While, the TDS levels in the other two wells will decrease in comparison with the previous five-year period. Our results contribute in giving more timely information for better water quality management and demand water planning. It also showed the importance of the selected models in TDS forecasting especially the increase in TDS levels in the most of the examined wells in this study. Consequently, water authorities have to use these wells wisely to decrease the effects of TDS on the population.

For future work, more forecasting models can be used in addition of applying machine learning and deep learning techniques in TDS forecasting utilizing additional wells' characteristics and features. Although the TDS forecasting is limited to the wells of Khan Younis city, other areas of groundwater as well as data collected for more wells can be also examined.

References

[1] Abduljaleel, H., Sameen, M., Ahmed, A., Azzam, R., & Schuttrumpf, H. (2018). Prediction and forecasting of total dissolved solids (TDS) by recurrent neural networks. Journal of Advanced Research in Dynamical and Control Systems, 10, 1922-1944.

[2] Garcia, J., Heo, J., & Kim, C. (2024). Machine Learning Algorithms for Water Quality Management Using Total Dissolved Solids (TDS) Data Analysis. Water, 16(18), 2639.

[3] Aiash, M., & Mogheir, Y. (2017). Comprehensive Solutions for the Water Crisis in Gaza Strip. IUG Journal for Natural Studies, 25(3).

[4] Alda, S., & Amorós, A. M. R. (2018). Water Crisis in Palestines ans its Affects on Agricultura Sector. In Congreso Nacional del Agua Orihuela: Innovación y Sostenibilidad (pp. 311-324). Universidad de Alicante/Universitat d'Alacant.

[5] Adhikari, R., & Agrawal, R. K. (2013). An introductory study on time series modeling and forecasting. arXiv preprint arXiv:1302.6613.

[6] Elsworth, S., & Güttel, S. (2020). Time series forecasting using LSTM networks: A symbolic approach. arXiv preprint arXiv:2003.05672.

[7] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. PloS one, 13(3), p.e0194889.

[8] Kim, K. J. (2003). Financial time series forecasting using support vector machines. Neurocomputing, 55(1-2), 307-319.

[9] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 1394-1401). Ieee.

[10] Sadorsky, P. (2021). A random forests approach to predicting clean energy stock prices. Journal of Risk and Financial Management, 14(2), 48.

[11] Gómez-Omella, M., Esnaola-Gonzalez, I., & Ferreiro, S. (2021, January). Short-Term Electric Demand Forecasting for the Residential Sector: Lessons Learned from the RESPOND H2020 Project. In Proceedings (Vol. 65, No. 1, p. 24). MDPI.

[12] Moein, E., & Awasthi, A. (2020). Carsharing customer demand forecasting using causal, time series and neural network methods: a case study. International Journal of Services and Operations Management, 35(1), 36-57.

[13] Liu, Z., Du, W., Yan, D. M., Chai, G., & Guo, J. H. (2018). Short-term traffic flow forecasting based on combination of k-nearest neighbor and support vector regression. Journal of Highway and Transportation Research and Development (English Edition), 12(1), 89-96.

[14] Bellotti, A., Brigo, D., Gambetti, P., & Vrins, F. (2021). Forecasting recovery rates on non-performing loans with machine learning. International Journal of Forecasting, 37(1), 428-444.

[15] Abbasimehr, H., & Shabani, M. (2019, March). Forecasting of customer behavior using time series analysis. In The 7th International Conference on Contemporary Issues in Data Science (pp. 188-201). Cham: Springer International Publishing.

[16] Lee, S., Lee, K. K., & Yoon, H. (2019). Using artificial neural network models for groundwater level forecasting and assessment of the relative impacts of influencing factors. Hydrogeology Journal, 27(2).

[17] Kouziokas, G. N., Chatzigeorgiou, A., & Perakis, K. (2018). Multilayer feed forward models in groundwater level forecasting using meteorological data in public management. Water resources management, 32(15), 5041-5052.

[18] Guzman, S. M., Paz, J. O., Tagert, M. L. M., & Mercer, A. E. (2019). Evaluation of seasonally classified inputs for the prediction of daily groundwater levels: NARX networks vs support vector machines. Environmental Modeling & Assessment, 24, 223-234.

[19] Nasr, M., & Zahran, H. F. (2014). Using of pH as a tool to predict salinity of groundwater for irrigation purpose using artificial neural network. The Egyptian Journal of Aquatic Research, 40(2), 111-115.

[20] Kim, M., Kim, Y., Kim, H., Piao, W., & Kim, C. (2016). Evaluation of the k-nearest neighbor method for forecasting the influent characteristics of wastewater treatment plant. Frontiers of Environmental Science & Engineering, 10, 299-310.

[21] El-Jaat, M., Hulley, M., & Tétreault, M. (2018). Evaluation of the fast orthogonal search method for forecasting chloride levels in the Deltona groundwater supply (Florida, USA). Hydrogeology journal, 26(6), 1809-1820.

[22] El Bilali, A., Taleb, A., & Brouziyne, Y. (2021). Groundwater quality forecasting using machine learning algorithms for irrigation purposes. Agricultural Water Management, 245, 106625.

[23] Zhang, Y., & Wang, J. (2016). K-nearest neighbors and a kernel density estimator for GEFCom2014 probabilistic wind power forecasting. International Journal of forecasting, 32(3), 1074-1080.

[24] Thu, T. H., Van Nghia, N., Huong, T. T. T., Van Thang, D., & Hien, N. T. (2017). Saltwater Intrusion Forecast of the Pleistocene Aquifer Caused by Groundwater Exploiting in the Nam Dinh Coastal Zone. In From Science to Society: New Trends in Environmental Informatics (pp.

83-92). Cham: Springer International Publishing.
[25] Aryafar, A., Khosravi, V., Zarepourfard, H., & Rooki, R. (2019). Evolving genetic programming and other AI-based models for estimating groundwater quality parameters of the Khezri plain, Eastern Iran. Environmental earth sciences, 78, 1-13.
[26] Abuelaish, B., & Olmedo, M. T. C. (2018). Analysis and modelling of groundwater salinity dynamics in the Gaza strip. Cuadernos geográficos de la Universidad de Granada, 57(2), 72-91.
[27] Amra, I. A. A., & Maghari, A. Y. (2018, October). Forecasting groundwater production and rain amounts using ARIMA-Hybrid ARIMA: Case study of Deir El-Balah City in GAZA. In 2018 International Conference on Promising Electronic Technologies (ICPET) (pp. 135-140). IEEE.
[28] Mohd Faizan Rizvi1, Shivang Sahu2, Dr. Sadhana Rana3, ARIMA Model Time Series Forecasting.
[29] A.C. Müller, S. Guido, Introduction to Machine Learning with Python: a Guide for Data Scientists, O'Reilly Media, Inc., 2016.
[30] Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 1394-1401). Ieee.
[31] Tofallis, C. (2015). A better measure of relative prediction accuracy for model selection and model estimation. Journal of the Operational Research Society, 66(8), 1352-1362.
[32] Abuamra, I. A., Maghari, A. Y., & Abushawish, H. F. (2021). Medium-term forecasts for salinity rates and groundwater levels. Modeling Earth Systems and Environment, 7(1), 485-494.
[33] Mushtaha, A. M., Van Camp, M., & Walraevens, K. (2019). Evolution of runoff and groundwater recharge in the Gaza Strip over the last four decades. Environmental Earth Sciences, 78(1), p.32.

Asem H. Shurrab holds a master degree in information technology from Islamic University of Gaza, Palestine. He is specialized in data science especially data mining and forecasting and he published many research papers in this field.

Abdelkareem M. Alashqar received his PhD in Information Systems from Mansoura University, Egypt in 2017. He is an assistant professor at the department of software development, Islamic University of Gaza, Palestine. His research and teaching interests include software engineering, software quality evaluation, natural language processing and machine/deep learning.

Ashraf Y. Maghari is an associate professor of computer science at Islamic University of Gaza. He holds a PhD in computer vision and image processing from USM university, Malaysia. He has extensive research experience in computer science fields such as data/text mining, image processing, computer vision and deep learning.