TRACKING AND FORECASTING HEAVY METAL WATER CONTAMINATION

[1]DVH Venu Kumar, [2]K.Bala Krishna

[1,2]Assistant Professor, [1,2]Department of Computer Science & Engineering, Geethanjali Institute of Science and Technology, Gangavaram, Andhra Pradesh, India

Abstract

Water is crucial for all types of life. The nature of water helps in controlling in the biotic diversity, variety, and rate of sucession . The disintegrating nature of common water assets like lakes, Streams, and estuaries is one of the direst and most troubling issues looked by humankind. The critical issue of water pollution caused by heavy metals, presenting a comprehensive approach to monitoring and forecasting. The escalating threat to ecosystems and human health necessitates innovative solutions. Our project integrates advanced sensing technologies (ARIMA), data analytics, and predictive modeling to create a real-time monitoring system. By leveraging these tools, we aim to not only detect and quantify heavy metal concentrations in water but also forecast potential pollution events. This proactive approach will contribute to timely interventions, safeguarding water resources and minimizing environmental impact. The project's outcomes hold promise for enhancing water quality management and ensuring a sustainable and healthier future. To assess the performance of the developed model, the metrics used are Mean Squared Error, Root Mean Squared Error, and Regression Analysis. This proactive approach will contribute to timely interventions, safeguarding water resources and minimizing environmental impact. The project's outcomes hold promise for enhancing water quality management and ensuring a sustainable and healthier future

INTRODUCTION OF WATER POLLUTION

Water assumes a crucial role in our day by day life and the nature of water in a region intensely influences the practical improvement of nearby ordinary industrial, agricultural and other anthropogenic activities. Common water resources like groundwater and surface water have dependably been the least expensive and most broadly accessible sources of freshwater. In any case, these assets are destined to progress toward becoming defiled because of different variables including human, industrial and commercial activities just as common procedures. Notwithstanding that, poor sanitation foundation and absence of mindfulness additionally contribute enormously to drinking water defilement. A considerable lot of the water pollutants have long haul negative effects on water quality, counseling a hazard to human wellbeing. Poor water quality influences the earth and human prosperity. Therefore, freshwater is seriously diminished. Additionally, contaminated water can prompt some waterborne ailments and furthermore impact child mortality. As indicated by the United Nations, waterborne infections cause the death of 1.5 million children for every year. The World Health Organization says that consistently more than 3.4 million individuals die because of water-related ailments. In this way, it is extremely essential to devise novel methodologies and techniques for deteriorating water quality and to figure future water quality patterns. So as to complete valuable and productive water quality analysis and foreseeing the water quality examples, it is important to incorporate a temporal dimension to the analysis, with the goal that the seasonal variation of water quality is tended to. Distinctive approaches have been proposed and applied for analysis and checking water quality and time series analysis.

PURPOSE

Predictive analysis can help to capture relationships among numerous variables that can help to assess risk with a particular set of conditions. The purpose is to propose an ARIMA considering better and increasingly precise data to foresee and assess the water quality. This will help in foreseeing the estimations of water quality parameters dependent on their present qualities.

PROBLEM STATEMENT

There is an assortment of strategies utilized for water quality prediction at home and abroad. These strategies are principally separated into four classifications:-

1. Mathematical Statistics
2. LSTM
3. Neural Networks
4. ARIMA model

The strategy for mathematical statistical is powerful in modelling; however, the prediction is not perfect. The strategy for the gray theory is only appropriate for approximating exponential functions, not for the total non-linear functions. The strategy for chaos theory simply can be valuable when the training data is extremely well-off. The traditional neural network, whose advantages are non-linearity, self-organization learning is suitable for managing non-linear, arbitrary data. In light of the structure of the traditional neural network, it is not appropriate for dealing with the time series data. Due to the few shortcomings from different models, LSTM model is picked so as to build up a far-reaching approach for effective water quality prediction and analysis. In spite of the fact that ARIMA models are very adaptable in speaking to various sorts of time series, AR, MA, and consolidated AR and MA(ARMA), their major impediment is the pre-assumed linear form of the model.

OVERVIEW OF DATA

Use of pesticides and fertilizers that contain metals like arsenic and cadmium can lead to runoff into nearby water bodies. Wastewater Disposal: Untreated or inadequately treated wastewater from domestic and industrial sources often contains heavy metals. Atmospheric Deposition: Heavy metals present in the atmosphere, emitted by industrial and vehicular emissions, can deposit into bodies of water. Natural Sources: Natural geological processes can release heavy metals into water. For example, arsenic is commonly found in high concentrations in groundwater in some regions due to the local geology. Impacts of Heavy Metal Pollution Human Health: Exposure to heavy metals can lead to a range of health issues, including but not limited to neurological damage, renal failure, cancer, and developmental problems in children. For example, mercury poisoning can cause neurological and behavioral disorders. Aquatic Life: Heavy metals can accumulate in the tissues of aquatic organisms, leading to poisoning, which can disrupt aquatic ecosystems. Bioaccumulation and biomagnification can cause elevated levels of toxicity in the food chain, affecting species up the trophic levels. Environmental Degradation: Heavy metal contamination can lead to reduced biodiversity and the disruption of natural processes in ecosystems, impacting the quality of the soil and water and thereby affecting plant and animal life.

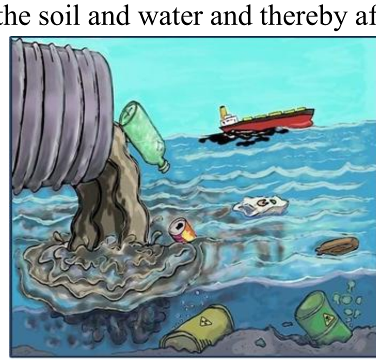

Fig.1. polluted water by heavy metal

We’re looking to delve deeper into forecasting heavy metal water contamination, consider refining your approach with a specific title like "Predictive Modeling of Heavy Metal Water Contamination: Integrating Data Science and Environmental Monitoring." This title hints at the comprehensive approach you're taking, combining predictive modeling techniques with real-time environmental monitoring to anticipate and mitigate heavy metal pollution in water sources

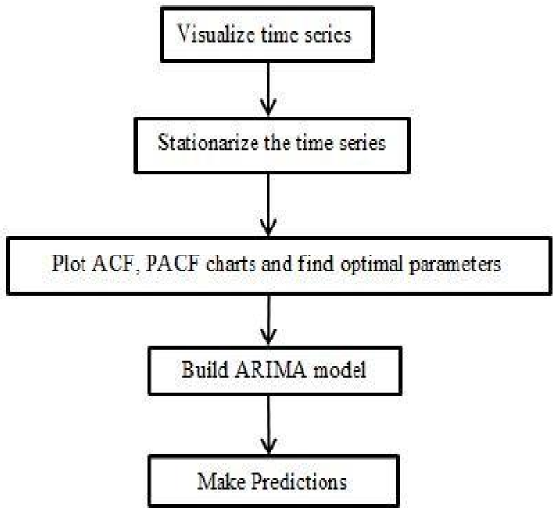

Fig.2. Model workflow

OBJECTIVE

The fundamental thought of this research is to devise a comprehensive methodology that analyzes and predicts the water nature of specific regions with the assistance of certain water quality parameters. These parameters incorporate physical, biological and chemical factors which impact water quality. This research expects to address this issue by recommending a model dependent on Machine Learning procedures so as to anticipate the future water quality patterns of a specific region with the assistance of historical water quality data. ARIMA model is utilized to build up a methodology for viable water quality forecast and analysis. The model based on ARIMA for water quality forecast is contrasted with two models: one with the traditional neural network and other based on ARIMA model. The results verify the efficiency of the model we have proposed

Literature Review

Applying and comparing LSTM and ARIMA to predict CO levels for a time-series measurements in a port area

A detailed description of the LSTM and an application regarding field data forecasting can be found in. In addition, the original LSTM is very slightly not enough in explaining the relationship of the input and output of the network, to solve this problem, the attention-based mechanism is inserted into LSTM. Moreover, in local air pollution measurements close to main streets, which reside on small Internet of Things (IoT) devices.

we collect data from a set of environmental sensors, which form a wireless environmental sensor located in the broader area of the port of Igoumenitsa in Greece. The station measures the most important pollutants including the CO in the port. We analyse the CO measurements at 6-h intervals for the entire dataset. We show that there exists an abnormal rise in values. Thereafter, we transform the dataset to a moving average values dataset, in order to reduce extremely high values.

The moving average is a popular method by which to transform a dataset to a smoother problem, in terms of the values used for the prediction phase. Moving average smoothing is a naive and effective technique in time series forecasting. It can be used for data preparation, feature engineering, and even directly for making predictions. Importantly, we utilise machine learning and especially LSTM to show the predicted outcome of the collected values from the port by using different batches of the procedure.

Forecasting water quality parameters using artificial neural network for irrigation purposes

Water pollution has posed a major problem and identifying the points of pollution in the River system is a very difficult task. To overcome this task, the need to determine the pollution level arose by modeling and predicting four water quality parameters at four (4) different locations using the Artificial Neural Network

The testing model performance shows that the R^2 value ranges from 0.952 to 0.967, 0.953 to 0.970, 0.951 to 0.967and 0.953 to 0.968, for pH, TDS, EC and Na while the forecast performance evaluation shows that the R^2 values ranges from 0.945 to 0.968, 0.946 to 0.968, 0.944 to 0.967 and 0.949 to 0.965 for pH, TDS, EC and Na respectively. It was also observed that the Root Mean Squared Error (RMSE) ranges from 0.022 to 0.088, 0.012 to 0.087, 0.015 to 0.085 and 0.014 to 0.084 for pH, TDS, EC and Na, respectively. Information from this study will serve as a guide to researchers on the water quality index for irrigation purposes. Also, it will guide the government and agencies on policy, management and decision-making on water resources.

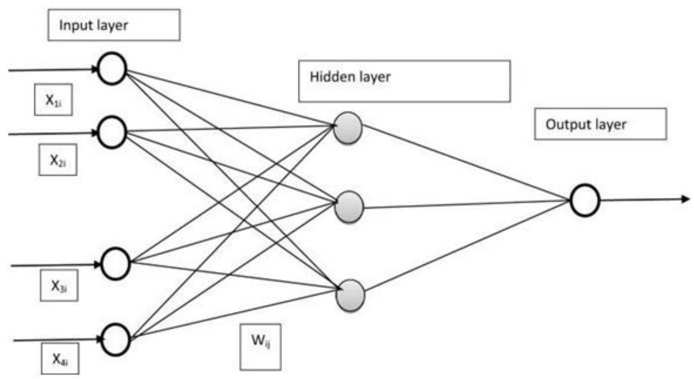

Fig.3. artificial neural networks

Artificial neural network architecture

This network architecture contains a number of layers namely; the input, hidden and output layers, respectively. The architecture determines the number of connection weights and also the way information flows through the network. The determination of the best network architecture is one of the difficult tasks in the artificial neural network model building process but one of the most important steps that must be taken. In this study, the river quality parameters

at times t_{-n} was used to predict the river quality parameters at time t_{n+1}. The Neuron class describes an entity with an (x, y) location that manages an array list of neurons, as well as its own location that are drawn relative to the network's center. and so forth. Consequently, high layers of the SAE tend to learn higher- order features and have very good discriminative power.

Ground water quality forecasting modelling using Artificial intelligence

This review paper closely explores the techniques and significances of the most potent artificial intelligence (AI) approaches in a concise and integrated way, specifically in the groundwater quality modelling and forecasting for its suitability in domestic usage. This paper systematically provides an extensive review of the four most used AI methods: artificial neural network (ANN), adaptive network-based fuzzy inference system (ANFIS), evolutionary algorithm (EA) and support vector machine (SVM), to reflect on the features.

The efficiently managing and improving groundwater quality are the accurate and reliable forecast of the groundwater resource information, where better understanding of the hydrogeological process and behaviour is fundamental. However, the complex and changeable characteristics of the hydrogeological system require specific tools to develop such predictions. Among the various approaches, groundwater modelling has proven to be a convenient tool to understand the groundwater systems and identify the current groundwater threats (1978).

Air quality forecasting using artificial neural networks with real time dynamic error correction in highly

polluted regions

Air pollution is an important issue, especially in megacities across the world. There are emission sources within and also in the regions around these cities, which cause fluctuations in air quality based on prevailing meteorological conditions. Short term air quality forecasting is used not to just possibly mitigate forthcoming high air pollution episodes, but also to plan for reduced exposures of residents. In this study, a model using Artificial Neural Networks (ANN) has been developed to forecast pollutant concentration of PM_{10}, $PM_{2.5}$, NO_2, and O_3 for the current day and subsequent.

Total concentrations and sources of heavy metal pollution in global river and lake water bodies

This study collected past sampling data on total concentrations of 12 heavy metals (Cd, Pb, Cr, Hg, Zn, Cu, Ni, Al, Fe, Mn, As, and Co) in surface water bodies, i.e., 168 rivers and 71 lakes, from 1972 to 2017. The intent was to investigate the levels and sources of heavy metal pollution across five decades and five continents. Mean heavy metal concentrations in global river and lake water, and the number of heavy metals with concentrations greater than the published threshold limits as per the standards of both the World Health Organization (WHO) .

PCA was used to represent the total variability of the original metal data in a minimum number of factors. Each factor is orthogonal to all others, which results in the smallest possible covariance. The first factor represents the weighted (factor loading) linear combination of the original variables that account for the greatest variability.

EXISTING SYSTEMS

Various analytical models are used to predict water quality, which can be divided into traditional, based on statistical. models, and non-traditional, using artificial intelligence (AI) approaches. Methods of KNN, machine learning, artificial neural networks (ANN) can be attributed to non-traditional methods, and regression analysis, time series methods can be attributed to traditional methods . A review of the literature has shown that artificial neural networks (ANN) are at the peak of popularity in modeling the prediction of water pollution

PROPOSED SYSTEM

ARIMA

ARIMA is an acronym for "Auto Regressive Integrated Moving Average ".Any 'non- seasonal' time series that exhibits patterns and is not a random white noise can be modeled with ARIMA models. Time series forecasting are predicting future values over the period of time. you use only the previous values of the time series to predict its future values, it is called Univariate Time Series Forecasting.

use predictors other than the series (a.k.a exogenous variables) to forecast it is called Multi Variate Time Series Forecasting.

1. AUTO REGRESSIVE(AR) COMPONENT
2. INTEGRATED (I) COMPONENT
3. MOVING AVERAGE(MA) COMPONENT

ARIMA model order is depicted as (p,d,q) with a number of times the function occurs in running the model.

p - number of lag observations d - degree of differencing.

q - size of the moving average window

The AR part of ARIMA shows that the advancing variable of interest is relapsed on its lagged values. The MA part demonstrates that the regression error is really a direct combination of the error terms whose values came contemporaneously and at different times in the past. The I shows that the data values have been supplanted with the difference between their values and the past values. This differencing procedure may have been executed more than once. The reason for each one of these features is to make the model fit the information just as conceivable.

AUTO REGRESSIVE AR

It refers to a model that shows a changing variable that regresses on its own lagged, or prior, values. ARIMA model order is depicted as (p,d,q) with a number of times the function occurs in running the model. The 'p' parameter represents the number of lag observations included in the model.
The autoregressive (AR) part of the model is explained by the following equation:

$$= + \emptyset 1-1 + \emptyset 2-2 + \cdots + \emptyset - +$$

INTEGRATED
It represents the differencing of raw observations to allow the time series to become stationary (i.e., data values are replaced by the difference between the data values and the previous values).
Non-seasonal ARIMA models are usually indicated by ARIMA (p,d,q) where parameter p - number of lag observations added in the model, also known as the lag order.
d - number of times that the raw observations are differenced, also known as the degree of differencing.
q - size of the moving average window, also known as the order of moving average.

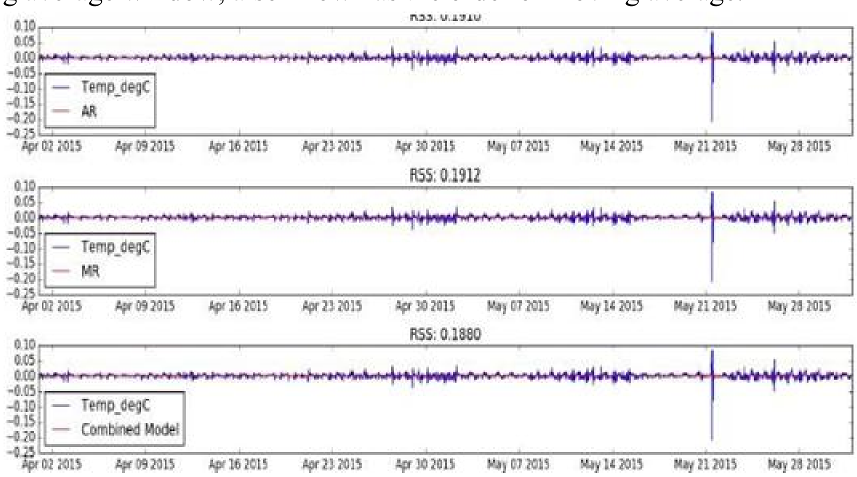

Fig.4. AR, MA and Combined Model for Temperature (April – May 2021)

MOVING AVERAGE
It incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations. The MA part demonstrates that the regression error is really a direct combination of the error terms whose values came contemporaneously and at different times in the past.
The moving average (MA) part of the model is explained by the following equation:

$$= + \emptyset_{1-1} + \emptyset_{-} + \cdots + \emptyset_{-} +$$

System Architecture
The architecture of the proposed decision fusion is based on the time series forecasting as shown. The time series forecasting models chosen are ARIMA. The proposedsystem is implemented in two phases i.e., with pre-trained and fine- tuned time series models. In the pre-trained models implementation, regularization is not applied and the pre-trained future values are used and for the fine-tuned implementation, regularization is applied to predict the future concentration levels.

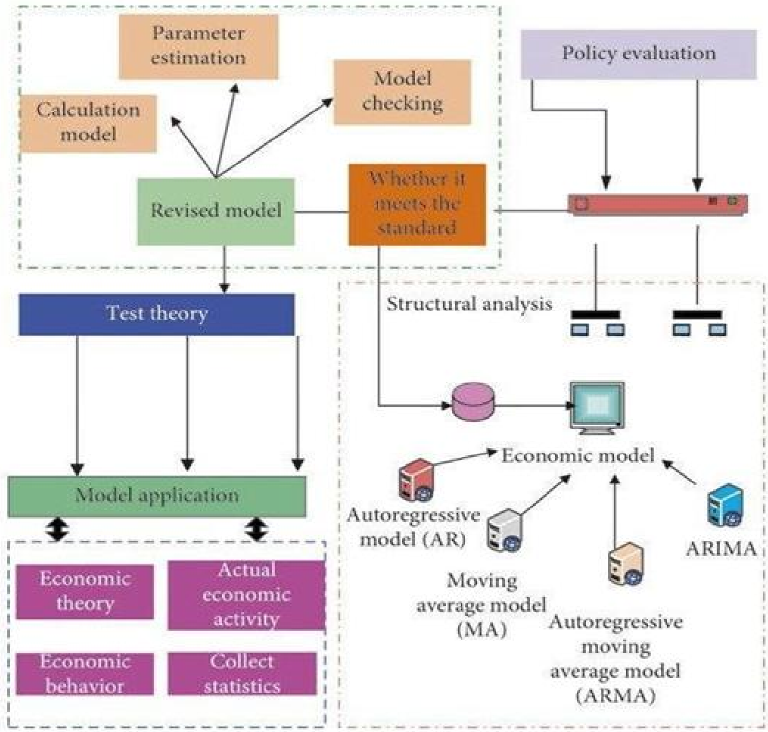

Fig.5. ARIMA Architecture

Model application

The application of an ARIMA model in time series analysis involves several key steps, from identifying the model parameters to forecasting future values. This process can be used across various fields such as finance, economics, manufacturing, environmental science, and more.

Structure analysis

Structured analysis in the context of ARIMA modeling involves systematic steps to prepare, identify, estimate, and validate an ARIMA model suitable for time series forecasting. Let's explore a more detailed step-by-step approach using an example of monthly sales data.

Policy evaluation

An ARIMA (AutoRegressive Integrated Moving Average) model typically involves analyzing how a particular policy change or event impacts a time series data set. ARIMA models are highly effective in forecasting and can be utilized to understand and measure the effects of policy changes on a given metric over time.

Data Exploration: Understand the time series data you're working with. Explore its patterns, trends, and seasonality.

Stationarity: Check if the time series is stationary. If not, apply differencing until it becomes stationary. The number of differences needed determines the 'd' parameter in ARIMA(p, d, q).

ACF and PACF Plots: Examine the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots to identify potential values for 'p' and 'q' respectively. These plots help determine the orders of the autoregressive and moving average components.

Model Selection: Based on the ACF and PACF plots, select initial values for 'p' and 'q'. You might also try different combinations and compare model performance using techniques like AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion).

Model Estimation: Use software like Python's statistical models or R to estimate the parameters of the ARIMA model.

Model Evaluation: Evaluate the model's performance using diagnostic tests, such as examining residuals for autocorrelation and normality.

Forecasting: Once the model is validated, use it to make forecasts. Monitor the forecasts against actual values and adjust the model as needed

Refinement: Refine the model by iterating through steps 3 to 7, adjusting parameters and evaluating performance until you achieve a satisfactory model.

DATASET

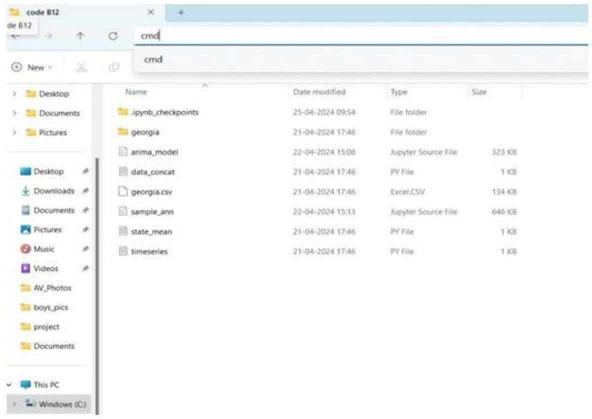

Fig.6. Document of all files

[4]:

	date	pH
0	2014-10-01	7.009447
1	2014-10-02	7.044741
2	2014-10-03	6.972112
3	2014-10-04	6.928980
4	2014-10-05	6.980445
5	2014-10-06	7.019258
6	2014-10-07	7.031518
7	2014-10-08	7.031793
8	2014-10-09	7.013097
9	2014-10-10	7.007799

Fig.7. Dataset containing of water concentration levels

MODULES AND ITS IMPLEMENTATION

DATA CLEANING AND PREPARATION

The data that we have collected from the USGS NWIS website, has been converted from a tab- separated values (.tsv) file format to a comma-separated values (.csv) file format. The data will be stored in a dataframe and date column has been set as an index. The rows with NaN values will be dropped from the dataframe. The data has been resampled on a daily basis and the daily average has been calculated. Then, the resampled data will be saved in another CSV file.

DATA PREPROCESSING

We load our information in an appropriate place and set it up to use in our machine learning training. This is likewise a decent time to do any appropriate visualizations of our information, to help check whether there are any pertinent connection between various factors we can exploit, just as to indicate us if there are any imbalances. The dataset gathered from USGS's online information archive initially contaied 1,50,729 rows.

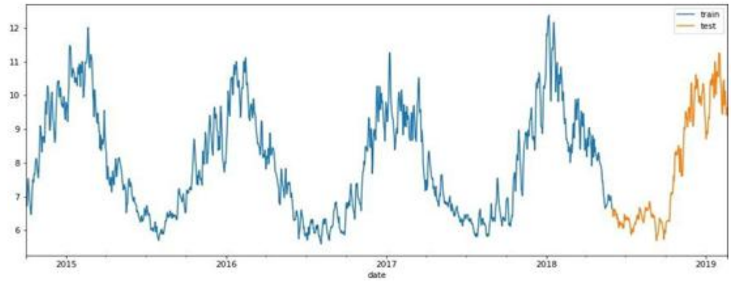

Fig.8.Trained data set of the sample

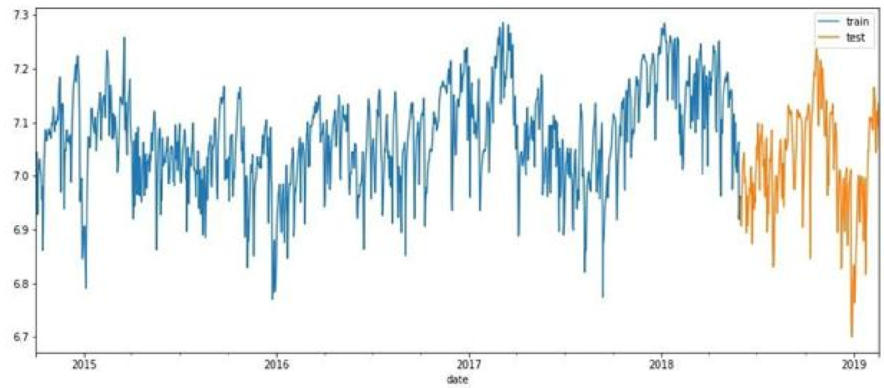

Fig.9. Trained and test data sample

Results

Step1: Open the file manager and move to the local disk(:c) and click on we have the MAJOR PROJECT FOLDER click on we have this code B12 FOLDER

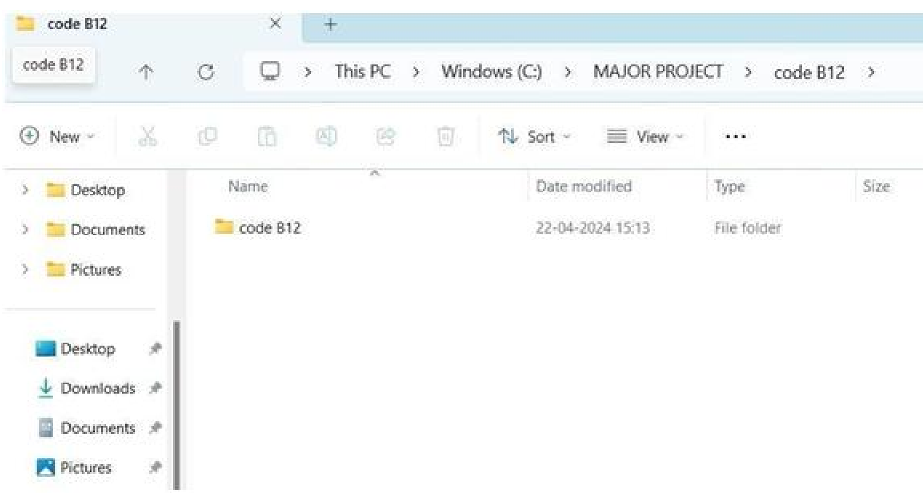

Fig.10. codeB12FOLDER

Step2: ClickonthecodeB12folderwehavetwofilesonezipfieanotherisinstalledfile. Click on code B12 we get the some files we have to redirect into command prompt.

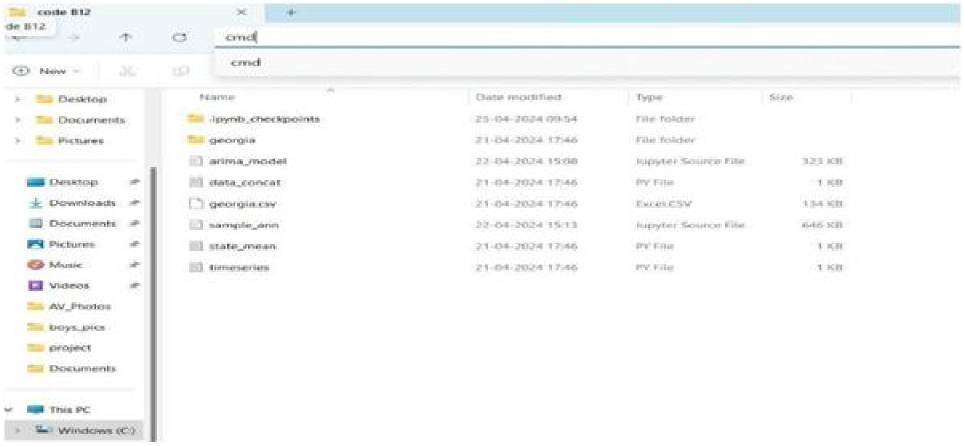

Fig.11. ARIMAData Files

Step 3 : In command prompt we have to enter the JUPYTER NOTEBOOK then redirect to the Chrome for the having multiple file open the **simple_ann file** run in the jupyter notebook.
We have the neural network prediction and forecast graph.

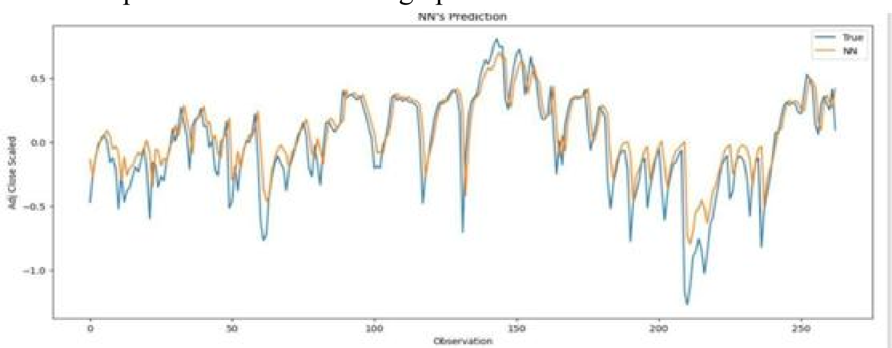

Fig.12. NEURAL NETWORK PREDICTION

Mean Square error for the neural network is 0.0437

Run the LSTM PREDICION which we get the forecasting graph.

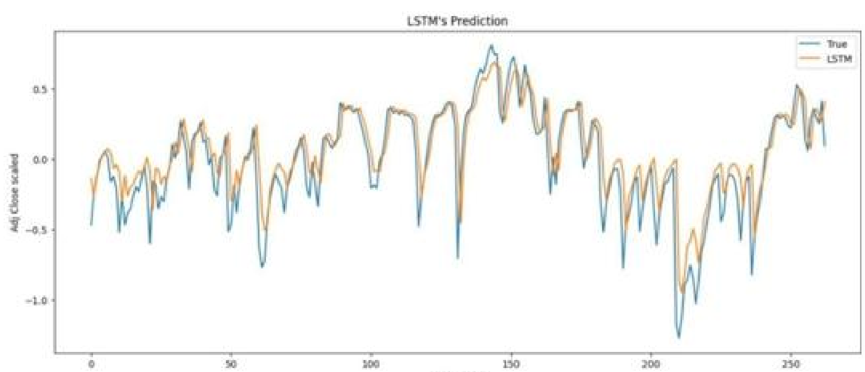

Fig.13. LSTM PREDICATION

Mean square error for the LSTM is 0.0414
From this LSTM is less error rate than Neural network . LSTM is the better performance

Step 4: Open the arima_model file run the model we get the ACF plots , PACF plots.

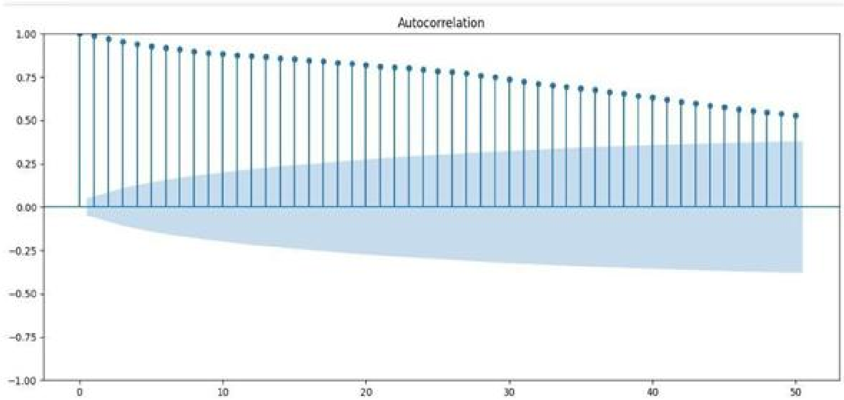

Fig.14. ACF PLOTS

We are construct the PACF plots by the model then we are constructed the residual error graph.

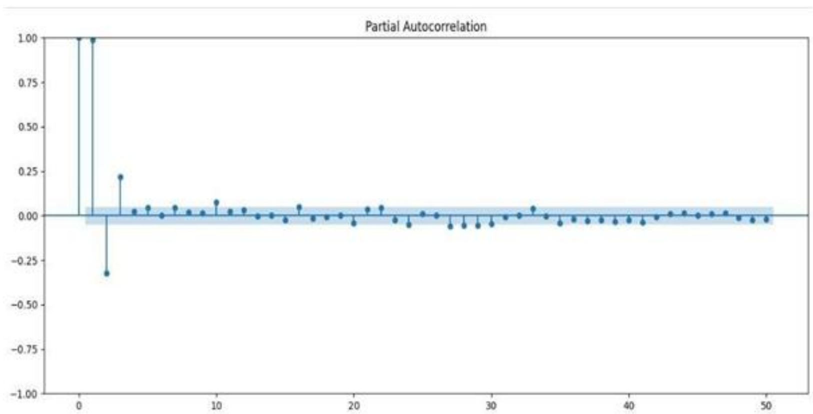

Fig.15. PACFPOLTS

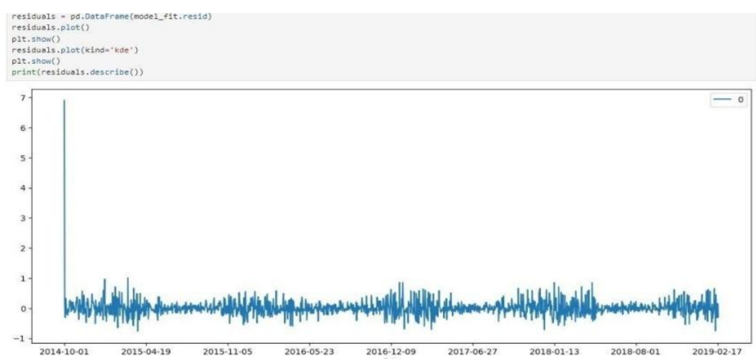

Fig.16. residual error plots

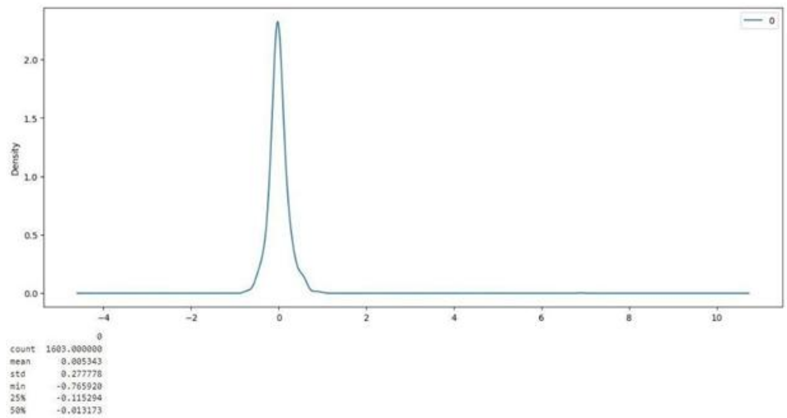

Fig.17. density vs concentration level

Step 5: Run the Arima model by using of the ACF and PACF plots to forecast the model.

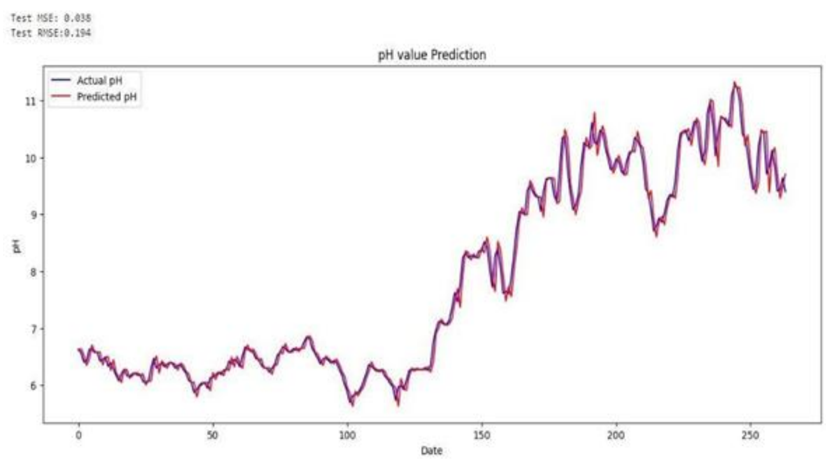

Fig.18. Arima forecasting

MODEL	R^2	MSE	RMSE
ARIMA	0.725	0.0380	0.194
LSTM	0.719	0.0414	0.2004
NN	0.696	0.0437	0.2262

Table.1. ARIMA, LSTM, NN error value

CONCLUSION

Contrasted with the famous forecasting models ANN ,LSTM and ARIMA, the prescient exactness of ARIMALSTM NN is higher which is plotted by graphs and depictedinperformancemetricsintermsofRMSE,MSE.Inaddition,ARIMAisless mean squares error and less root mean squares error. The experiment shows that the RMSE for the ARIMA is 0.194 and MSE is 0.038 which is better than other two famous forecasting models. The prediction results are stable. By this water contamination quality of Heavy metals can be choose this ARIMA model.

FUTURE SCOPE

WhenconsideringfutureworkforanARIMAmodelinthecontextofaprojectrelatedtoheavy metal water pollution, there are several directions and enhancements you can explore to improve understanding, forecasting, and management of pollution levels. Here are some focused areas where the ARIMA model could be expanded or refined for better utility in environmental science and policy planning:

Integration with Hydrological Models:

Coupling ARIMA models with hydrological models that predict water flow rates and accumulations could refine the predictions of heavy metal concentrations. This integrated modeling approach would take into account the dilution and concentration effects caused by varying water flows.

Advanced Computational Techniques:

Exploring machine learning techniques and combining them with traditional ARIMA models to handle non-linearities and complex interactions more effectively in pollution data.

Real-time Data Utilization and Forecasting:

Implementing the model with real-time data acquisition systems, such as IoT sensors in water bodies, can allow for dynamic updating and forecasting. This approach can enable more timely interventions and ongoing assessment of pollution control measures.

Exogenous Variables (ARIMAXorSARIMAX):

Adding exogenous variables to the ARIMA model (thus creating an ARIMAX or SARIMAX model) can provide deeper insights. Variables such as pH levels, water temperature, precipitation data, and industrial activity metrics can be crucial in predicting heavy metal concentrations more accurately.

References

[1] F. Thomas, Rural WaterSupplies and Water-Quality/HealthyHousing ReferenceManual. USA: National Center for Environmental Health, 2005, ch. 8. [Online]. Available: https://www.cdc.gov/nceh/publications/books/housing/06_hhm_final_chapter_08.pdf

[2] R. S. Bryan Swistock Clemens and E. William Sharpe, ‘‘Drinking water quality in rural Pennsylvania and theeffect of management practices/school of forest resources and institutes of energy and the environment,’’ Pennsylvania State Univ., State College, PA, USA, Tech. Rep., Oct. 2007.

[3] WaterResourcesofKazakhstanintheNewMillennium.Review(APublicationinSupport of the UN Millennium Development Goals. Goal 7: Ensure Environmental Sustainability), United Nations Develop. Program Kazakhstan, Almaty, Kazakhstan, 2004, p. 132. [Online].

[4] K. L. Valentini, E. E. Orolbaev, and A. K. Abylgazieva, Water Problems of Central Asia, Bishkek, Kyrgyzstan, 2004, p. 142.

[5] J. Hornero, M. Manzano, L. Ortega, and E. Custodio, ‘‘Integrating soil water and tracer balances, numerical modelling and GIS tools to estimate regional groundwater recharge:),’’ Sci. Total Environ., vol. 568, pp. 415–432, Oct. 2016, doi: 10.1016/j.scitotenv.2016.06.011.

[6] W. P. Kelley, ‘‘Permissible composition and concentration of irrigation water,’’ in Proc. ASCE, vol. 66, May 1940, pp. 607–613.

[7] B. K. Handa, ‘‘Modified classification procedure fro rating irrigation waters,’’ Soil Sci., vol. 98, no. 4, pp. 264–269, 1964.

[8] Environmental Protection in the Republic of Kazakhstan 15–19 Statistical Collection, Dept.Prod.Environ.Statist.,CommitteeStatist.MinistryNat.EconomyRepublicKazakhstan, Astana, Kazakhstan, 2019.

[9] J.I.Ubah,L.C.Orakwe,K.N.Ogbu,J.I.Awu,I.E.Ahaneku,andE.C.Chukwuma,

‘‘Forecastingwaterqualityparametersusingartificialneuralnetworkforirrigationpurposes,’’ Sci. Rep., vol. 11, no. 1, p. 24438, Dec. 2021, doi: 10.1038/s41598-021-04062-5.

[10] N. F. Che Nordin, N. S. Mohd, S. Koting, Z. Ismail, M. Sherif, and A. El-Shafie, ‘‘Groundwater quality forecasting modelling using artificial intelligence: A review,’’ Groundwater Sustain. Develop., vol. 14, Aug. 2021, Art. no. 100643, doi: 10.1016/j.gsd.2021.100643.

[11] S.V.Volkov,‘‘TechnologicalaspectsofwaterdisinfectionbyUVradiationText,’’ Technique,vol. 1,no. 12, pp.7–10, May 19

[12] J.A.Guzman,A.Shirmohammadi,A.M.Sadeghi,X.Wang,M.L.Chu,M.K.Jha,P.B. Parajuli, R. D. Harmel, Y. P. Khare, and J. E. Hernandez, ‘‘Uncertainty considerations in calibrationandvalidationofhydrologicandwater qualitymodels,’’Trans.ASABE,58,no.6, pp. 1745–1762, 2015, doi: 10.13031/trans.58.10710.

[13] E. D. Spyrou, I. Tsoulos, and C. Stylios, ‘‘Applying and comparing LSTM and ARIMA topredict COlevels for a time-series measurements ina portarea,’’Signals, vol.3,no. 2,pp. 235–248, Apr. 2022.

[14] S. Agarwal, S. Sharma, M. H. Rahman, S. Vranckx, B. Maiheu, L. Blyth, S. Janssen, P. Gargava, V.K. Shukla,and S.Batra, ‘‘Airquality forecastingusing artificial neural networks with real time dynamic error correction in highly polluted regions,’’ Sci. Total Environ., vol. 735, Sep. 2020, Art. no. 139454.

[15] K.Zhang,J.Thé,G.Xie,andH.Yu,‘‘Multi-stepaheadforecastingofregionalairquality using spatial–temporal deep neural networks: A case study of huaihai economic zone,’’ J. Cleaner Prod., vol. 277, Dec. 2020, Art. no. 123231.

[16] E. Eslami, A. K. Salman, Y. Choi, A. Sayeed, and Y. Lops, ‘‘A data ensemble approach for real-time air quality forecasting using extremely randomized trees and deep neural networks,’’ Neural Comput. Appl., vol. 32, no. 11, pp. 7563–7579, Jun. 2020.