Water Level Prediction of Lake: A Survey

Anjali Ramani[1] Apurv Verma[2]
[1]M. Tech. Scholar [2]Assistant Professor
[1,2]Department of Computer Science & Engineering
[1,2]MATS University, Raipur, India

Abstract— As a complex mathematical problem in water resources engineering and planning, water level variations in a lake represent the balance of water components for instance straight rainfall or evaporation at a lake's surface, groundwater exchange, and incoming or outgoing river flows. Lake water elevation changes seasonally example that, high in the soaked summer and low in the dried-out winter with sharp falling/ rising limbs through typhoon events, but not in a straightforward periodic mode except for the seiche oscillation that occurs mostly in huge lakes. Effectual forecasting tools play a significant role in the studies of lakes level measures. In this paper we have compared different literature.

Keywords: ARIMA, k-means, Genetic

I. INTRODUCTION

Forecasting is a procedure of estimating or predicting the future depends on past and nearby data. Forecasting provides information about the impending future measures and their consequences for the administration. It may not decrease the difficulties and hesitation of the future. Nevertheless, it increases the self-reliance of the management to craft imperative decisions. Forecasting is the foundation of premising. Forecasting uses various statistical data. Consequently, it is also called as Statistical Analysis.

Significance of forecasting involves following points:

- Forecasting provides reliable and relevant information about the present and past events and the probable future measures. This is very essential for sound planning.
- It gives self-belief to the managers for making imperative decisions.
- It is the source for making planning grounds.
- It keeps managers alert and active to face the challenges of future measures and the changes in the atmosphere.

Boundaries of forecasting involves following points:-

- The analysis and collection of data about the present, history and future involves lots of time and capital. Consequently, managers have to equilibrium the cost of forecasting with its reimbursement. Most of the small firms don't do forecasting on account of the high cost.
- Forecasting task can only approximate the future measures. It cannot pledge that these measures will take place in the future. Long-term prediction will be fewer accurate in comparison with to short-range forecast.
- Data Prediction is based on convinced assumptions. If these assumptions are mistaken, the forecasting will be incorrect. Forecasting is depend on past measures. On the other hand, past may not reiterate itself at all times.
- Forecasting need proper skills and judgement on the part of managers. Forecasts may go incorrect due to terrible judgement and skills on the part of some of the managers. Consequently, predicting data are subject to human error.

Lake water elevation changes seasonally example that, high in the soaked summer and low in the dried out winter with sharp falling/ rising limbs through typhoon events, but not in a straightforward periodic mode except for the seiche oscillation that occurs mostly in huge lakes. Effectual forecasting tools play a significant role in the studies of lakes level measures. These can be used to replicate the lake water elevation diversity based upon the presented measured data and forecast the probable responses under dissimilar scenarios, supporting management decisions of valuable water capital.

Forecasting of water elevations in the lakes [Nwobi-Okoye et. Al. 2013] in cities that reaches to city people for expenditure is mostly done physically and has the restraint that it cannot take into deliberation the past levels of water accessibility for forecasting the future levels patterns that can help in better planning and can shun the shortage of the water in Cities to a greater coverage. Forecasting approaches can be classified into various approaches –

A. Qualitative Approach

In this approach there is no use of any mathematical model due to the fact that the data available is not considered to be contributing to the future values (long-term forecasting)

B. Quantitative Approach

In this approach the historical data are available. It is based on analysis of historical data having the time series [H. Aksoy 2013] of particular variable and other related time series. It also examines the cause-and-effect relationships of one type of variable vs. other relevant variables

C. Time Series Approach

In this approach we have a single variable that keeps changing with time and whose future values are definitely related in some form to its past values.

The availability and accessibility of near real-time water level data at multiple locations for each of the lakes was a key component to the success of the operational forecasting system. The database of historical water levels and wave heights were used to calibrate the hydrodynamic and wave models for each lake, while the near real-time water level data was used for data assimilation on the hydrodynamic models.

The ease of use and convenience of near real-time water level data at manifold locations for every of the lakes was a key constituent for the accomplishment of the prepared forecasting system. The dataset of chronological water levels and wave heights were worn to standardize the hydrodynamic and wave models for apiece lake, while the near real-time water level data was used for data absorption on the hydrodynamic models.

Correct prediction of water level vacillation is significant in lake management due to its noteworthy impacts in an assortment of aspects. This attempt to examine and suggest a computer platform which can be used in enduring

for reservoirs level statement comes very useful. It's recognized that the system is with no trouble scalable to include supplementary knowledge for different reservoirs and always perform professionally.

Forecasting of water lake level can be useful for:-

- Assessing the severity of drought
- Predicting where/when there is a risk of flooding
- Limiting the impact of flooding and drought by enabling government and other agencies to put emergency response plans into operation

Identifying sensitive and risk prone areas for flooding and drought throughout the province Forecast is merely a prediction about the future values of data. However, most extrapolative model forecasts assume that the past is a proxy for the future. There are many traditional models for forecasting: exponential smoothing, regression, time series, and composite model forecasts, often involving expert forecasts. Regression analysis is a statistical technique to analyze quantitative data to estimate model parameters and make forecasts.

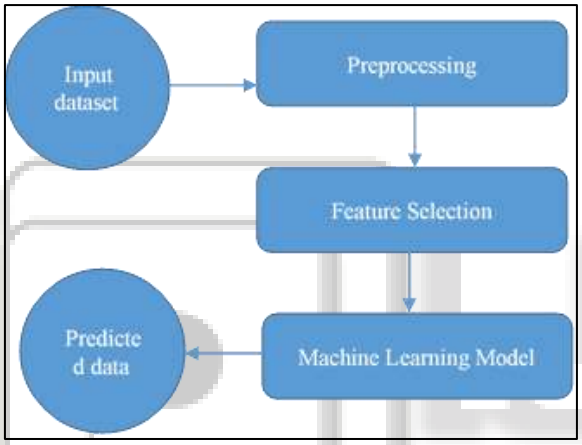

Fig. 1: Basic steps of data prediction

II. BACKGROUND STUDY

Mohammad Ali Ghorbani et. Al. a Multilayer Perceptron (MLP) based forecast model integrated with the Firefly Algorithm (FFA) as an optimizer tool has been adopted for the forecasting of the hydrological time series data: i.e., lake water level. The case study has been performed on a semi-arid region located at Lake Egirdir, in Turkey. By applying Average Mutual Information on the historical time-series of the lake water level data, a set of four input combinations of lake water level with lagged data series were considered to be most appropriate for the prediction of the 1-month lead time lake water level series. Several forecasting models have been developed, including the standalone MLP and the integrated MLP-FFA model using historical data for the period 1961–2016. The results have been evaluated with several statistical score metrics and visual displays; showing the better efficiency of the hybrid MLP-FFA [1 Springer 2017].

Tengfei Hu et. Al. developed an integrated modeling and analytical methodology for the water level management of lakes connected to Damregulated Rivers, and applies the methodology to the Dongting Lake in China. The following conclusions can be drawn:

1) The antecedent lake levels are the most important factor for the prediction of the current lake level;
2) The river discharge time lags selected by the GA well describe the spatial heterogeneity of the rivers' impacts on lake level changes;
3) The synchronized optimization is able to fulfill the potential of SVR, leading to highly accurate prediction of lake levels;
4) The integrated methodology can provide the lake level responses to future dam releases and the relative importance of different rivers in terms of affecting the lake level [2 Elesvier 2018].

Jalal Shiri et. Al. the Extreme Learning Machine (ELM) approach was used to predict the daily water-level in the Urmia Lake. Daily water-level data from the Urmia Lake in northwest of Iran were used to train, test and validate the employed models. Results showed that the ELM approach can accurately forecast the waterlevel in the Urmia Lake. Outcomes from the ELM model were also compared with those of genetic programming (GP) and artificial neural networks (ANNs). It was found that the ELM technique outperforms GP and ANN in predicting water-level in the Urmia Lake. It also can learn the relation between the water-level and its influential variables much faster than the GP and ANN. Overall, the results show that the ELM approach can be used to predict dynamics of water-level in lakes [3 Springer 2016].

Ozgur Kisi et. Al. [4] discussed that Forecasting lake level at various horizons is a critical issue in navigation, water resource planning and catchment management. In this article, multistep ahead predictive models of predicting daily lake levels for three prediction horizons were created. The models were developed using a novel method based on support vector machine (SVM) coupled with firefly algorithm (FA). The FA was applied to estimate the optimal SVM parameters. Daily water-level data from Urmia Lake in northwestern Iran were used to train, test and validate the used technique. The prediction results of the SVM–FA models were compared to the genetic programming (GP) and artificial neural networks (ANNs) models. The experimental results showed that an improvement in the predictive accuracy and capability of generalization can be achieved by the SVM–FA approach in comparison to the GP and ANN in 1 day ahead lake level forecast. Moreover, the findings indicated that the developed SVM–FA models can be used with confidence for further work on formulating a novel model of predictive strategy for lake level prediction [4 Ozgur Kisi 2015].

S.No.	Author/Year	Method Used	Description
1.	Prashant Shrivastava et. Al. Cloud Based Software Platform for Big Data Analytics In Water Reservoir Level Forecasting IOSR-JCE 2014 [5]	Autoregressive Integrated Moving Average (ARIMA) algorithmic	It stores the historic huge data set inside massive information storage and uses big data technologies to review ndianr and predict the future levels by applying data-driven analytics and data mining concepts.

2.	Chih-Chieh Young et. Al. Predicting the Water Level Fluctuation in an Alpine Lake Using Physically Based, Artificial Neural Network, and Time Series Forecasting Models Hindawi Publishing Corporation 2015 [3]	Artificial neural network (ANN) model (back propagation neural network, BPNN) & ARMAX	Four modelling approaches (the three-dimensional hydrodynamic model, ANN model, ARMAX model, and the combination model) have been implemented to predict water level fluctuation
3.	P. Delaney et. Al. Integration of real-time monitoring, modelling and Forecasting for the development and operation of the Great lakes storm surge operational system Montreal, Quebec, April 29 – May 2, 2015[4]	GLSSOS Method	GLSSOS has been in operation since 2007 and has proven to be a robust operational forecasting system capable of delivering accurate and reliable 48 hour forecasts of water levels and wave heights for the Great Lakes communities in Ontario. The level of accuracy demonstrated by the historical performance.
4.	K. H. V. Durga Rao et. Al. Kedarnath flash floods: a hydrological and hydraulic simulation study Elsevier Feb. 2014[22]	Hydraulic Simulation Study	The study reveals quantitative parameters of the disaster which was due to an integrated effect of high rainfall intensity, sudden breach of Chorabari Lake and very steep topography.
5.	C. J. Watras et. Al. Decadal oscillation of lakes and aquifers in the upper Great Lakes region of North America: Hydroclimatic implications AGU Publication 2014[5]	Unique hydrologic time series Analysis	Hydrologic analyses indicate that the oscillation has been governed primarily by changes in the net atmospheric flux of water (P-E) and stage-dependent outflow.
6.	Ozgur Kisi et. Al. A survey of water level fluctuation predicting in Urmia Lake using support vector machine with firefly algorithm Elsevier 2015[2]	SVM–FA models Used	The models were developed using a novel method based on support vector machine (SVM) coupled with firefly algorithm (FA). The FA was applied to estimate the optimal SVM parameters.
7.	Yong Liu et. Al. Quantitative evaluation of lake eutrophication responses under alternative water diversion scenarios: A water quality modeling based statistical analysis approach Elsevier 2013. [1]	Orthogonal test analysis and linear regression	A water quality modeling-based scenario analysis approach was applied to quantitatively evaluate the eutrophication responses
8.	Moslem Imani et. Al. Forecasting Caspian Sea level changes using satellite altimetry data (June 1992–December 2013) based on evolutionary support vector regression algorithms and gene expression programming Elsevier 2014 [17]	SVM	Accepted as an appropriate tool for modeling complex nonlinear phenomena in water bodies. In the study, they investigated the ability of two AI techniques: support vector machine (SVM), which is mathematically well-founded and provides new insights into function approximation, and gene expression programming (GEP), which is used to forecast Caspian Sea level anomalies
9.	Hakan Tongal et. Al. Phase-space reconstruction and self-exciting threshold modeling approach to forecast lake water levels Springer-Verlag Berlin Heidelberg 2013[19]	k-nearest ndianr (k-NN) model & SETAR model	A comparison of two nonlinear model approaches was made. Author used the k-NN approach and SETAR model for prediction of water levels for the three largest lakes in Sweden.
10.	T V Rajini kanth et. Al. Analysis of ndian weather data sets using data mining techniques CS & IT-CSCP 2014 [9]	J48 classification technique along with linear regression analysis	Author applied K-means cluster algorithm for grouping similar data sets together and also applied J48 classification technique along with linear regression analysis.
11.	Shubhendu Trivedi e. al. The Utility of Clustering in Prediction Tasks Centre for Mathematics and Cognition gran 2011 [16]	k-means	Observed that use of a predictor in conjunction with clustering improved the prediction accuracy in most datasets
12.	Tengfei Hu et. al./Water level management of lakes connected to	Genetic algorithm	This study develops an integrated modeling and analytical methodology for the water level

	regulated rivers: An integrated modeling and analytical methodology/ Elsevier 2018		management of lakes connected to dam regulated rivers, and applies the methodology to the Dongting Lake in China.

III. PROBLEM IDENTIFICATION

Lake water elevation prediction is extremely important for a truthful and dependable management of regional and local water resources. Lake water elevation modelling is significant for sustainable administration of such assorted objectives as hydropower generation, design of hydraulic structures and water supply.

Lake elevation changes are multifaceted outcomes of dissimilar site-specific meteorological variables and geographical, including precipitation, runoff, vanishing, temperature. The variables defining climate circumstances varies incessantly with time, prediction model can be urbanized either statistically or by using some different means like artificial neural networks, decision tree, regression, clustering techniques of data mining techniques. Weather prediction is a variety of data mining which is anxious with pronouncement of hidden patterns inside principally available meteorological data.

After studying different literature we found that if we apply clustering data mining technique then we can efficiently categories lake level historic data.

Traditional machine learning algorithm cannot give accurate result.

IV. DATASET USED

Accurate predictions of water level fluctuation that results from hydro meteorological variations and anthropogenic disturbances are needed for sustainable development and management of lake water usage.

Here to visualize the expected result of our project and how we can predict the lake water level by applying k-medoid clustering algorithm and multiple regression prediction model, we have demonstrated the process using Excel as follows:

Example Input Meteorological Data which is time series data taken from

URL: http://lakepowell.water-data.com/

CONTENT	INFLOW (cfs)	OUTFLOW (cfs)	HIGH TEMP	LOW TEMP	WATER TEMP	ELEVATION
12814028	9338	12755	0	0	54	3610.83
12821106	6109	12281	0	0	55	3610.9
12833245	9782	12651	47	32	55	3611.02
12839318	6391	12569	44	31	55	3611.08
12851469	9463	10679	40	26	55	3611.2
12854508	11109	9992	42	29	55	3611.23
12853495	6632	11032	50	32	55	3611.22
12862614	7118	9312	60	46	56	3611.31
12867682	3513	8469	54	32	56	3611.36
12877822	9502	11145	44	30	56	3611.46
12881880	5566	11079	52	34	56	3611.5

Fig. 2: Snippet of Dataset

V. CONCLUSION

In this project a portable and scalable prediction model is discussed for forecasting the water levels of lakes. Accurate predictions of water level fluctuation that results from hydro meteorological variations and anthropogenic disturbances are needed for sustainable development and management of lake water usage. Henceforth greater accuracy required for prediction system. Eventually we can conclude that if we use efficient machine learning algorithm.

REFERENCES

[1] Mohammad Ali Ghorbani1, Ravinesh C. Deo, Vahid Karimi, Zaher Mundher Yaseen, Ozlem Terzi, Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey, DOI 10.1007/s00477-017-1474-0 Springer 2017.

[2] Tengfei Hua,b, Jingqiao Maoa, Shunqi Panc, Lingquan Daid, Peipei Zhanga, Diandian Xua, Huichao Daia, Water level management of lakes connected to regulated rivers: An integrated modeling and analytical methodology Elsevier 2018.

[3] Jalal Shiri1 & Shahaboddin Shamshirband & Ozgur Kisi & Sepideh Karimi1 & Seyyed M Bateni & Seyed Hossein Hosseini Nezhad & Arsalan Hashemi, DOI 10.1007/s11269-016-1480 2016 Springer Science+Business Media Dordrecht 2016.

[4] Ozgur Kisi , Jalal Shiri , Sepideh Karimi , Shahaboddin Shamshirband , Shervin Motamedi , Dalibor Petkovic ´ f, Roslan Hashim Elsevier 2015 A survey of water level fluctuation predicting in Urmia Lake using support vector machine with firefly algorithm.

[5] Prashant Shrivastava1, S. Pandiaraj2 and Dr. J. Jagadeesan International Journal of Application or Innovation in Engineering & Management (IJAIEM) Volume 3, Issue 3, March 2014 Big Data Analytics In Forecasting Lakes Levels.

