Water Management Using Deep Neural Networks And Time Series Analysis

[1]Prof. Amit Uttarkar, [2]Akash Shukla, [3]Krushna Solanki, [4]Yash Shah, [5]Sonia Kamble

[1]Asst. Professor, [2]Student (B.E), [3]Student (B.E), [4]Student (B.E), [5]Student (B.E)

Department of Computer Engineering, JSPM Narhe Technical Campus, Pune.

Abstract **-** There are a lot of techniques that are already deployed at reservoir catchment stations, used to automate the process of discharge from the reservoir at mechanical levels, based on various reservoir metrics like depth and level. These are generally alarm systems that operate accordingly after occurrence of a certain threshold conditions of these mentioned parameters. This paper proposes an intelligent system to regulate the catchment processes so that these alarming conditions are avoided prior to their occurrence and ultimately avoiding any related malfunctioning. In our study we have made use of data from Poondi reservoir in Tiruvallur District of Tamil Nadu state to optimize the catchment flow by providing the decision makers with essentially important and relevant statistics. This paper makes use of advance deep learning techniques like neural network and time series analysis for ensuring more reliability of these generated statistics.

Index terms: Deep Neural Network, Time Series Analysis, Catchment Regulation.

I. INTRODUCTION

There have been a lot of machine learning techniques evolved with the time that can map to various real-time problems and transpose them into an efficient computer work flow. In the recent times it has become important to learn these techniques and use them appropriately for our benefits, as these can prove to be more sophisticated and reliable as compared to manual operations of the concerned processes. One of such area of current interest is water research, which can be strongly supported by various factors like increasing complexity of reservoir management due to urbanization, population growth, climate change. As these hydrological ecosystems are very essential for habituation of various living beings, it becomes extremely important to ensure proper supply and demand, and to prevent any malfunctioning of these water ecosystems.

One of the suitable ways of doing it could be by integrating these water flow mechanisms with various intelligent systems like high power computing networks. A key method to optimize the water flow can be prediction of various parameter with appropriate level of accuracy, that determine the flow indeed. A prior information of these parameter could better facilitate in various decisions to be taken. There are various such parameters and factors responsible for functioning of a reservoir like reservoir storage and rainfall recorded and inflow to the reservoir from the various river basins connected to the reservoir. Accurate information about rainfall can be crucial in reservoir operation and flooding prevention. But prediction of any kind of data in this context becomes very difficult due to a high complexity of the atmospheric processes accompanied with these datasets. There are various challenges in accurately predicting these data due to a non-linear relationship between these metrics, hence in our research we have made use of neural networks which learns about the non-linear relation between these datasets as well as between different values within themselves.

We present an intelligent system that does and provides important analytics essential for reservoir operation in form of one weak forecast of river inflow, rainfall and reservoir storage, warnings as required, and a suggestion about discharge volume based upon presented forecast and daily water usage of city. These forecasts are based upon an independent model that makes predictions exclusive of any other parameter and a dependent model that makes prediction dependent on other parameters. These predictions are made available on a daily basis as well as monthly basis so as to record different seasonal characteristics.

II. APPROACH

Most of the methods currently in use in the field of water management are based on the basic principle of Supply and Demand where the reservoir authorities make decision about the discharge in support to fulfill the demand made by the various consumers of it. This methodology was also one of the most traditional way of reservoir functioning, but with the time there were some glitches observed in this approach. We saw periods where water scarcity, floods were caused just because there were no informed decisions taken which led to reservoir drying and flooding. The reason for failure of these mechanism was the nonconsideration of various climate and environmental factors along with demands. This led to the revolution of water research where in people are trying to figure out various factors that may directly or indirectly affect the decisions about the catchment regulation and design systems that work around these factors, facilitating to make more informed decision. One of these researches was led by Loan Patri and his team [1] where they've designed a system that does real-time prediction and monitoring of river depth and rainfall at various nodes across the bank for optimizing the reservoir operations. The systems considered only the real-time data to make these prediction (forecast) rather than considering any historical trends and patterns in any of these factors.

We apply our analysis on Poondi reservoir in Tiruvallur District, Tamil Nadu. In our study we have considered three factors as essential for optimizing the water flow mechanism that are in order of relevance, inflow from various river basins, rainfall recorded particularly at the reservoir, and storage of the reservoir. Apart from using the real-time analysis to make some forecast for these values, we also considered any trend and seasonality in these factors. As studying historical patterns in these factors were equally important which can be understood by the fact that the values of these factors are very much affected by seasonal cycles, and in fact can prove to be effective in optimization. So eventually we get a more comprehensive water management process ensuring higher accuracy and sustainability of these hydrological ecosystems. And also, monsoon that plays a vital role in water management processes cannot be predicted using real-time data, it needs a seasonal decomposition. We have performed daily as well as monthly analysis which are followed by a module that comprises of some mathematical computations of discharge that aim to keep storage of reservoir at a constant value or evaluate the rate at which discharge should be incorporated in the coming days so as to avoid any vulnerability.

III. METHODOLOGY

We have designed the intelligent water optimization system using three layers (fig 1).

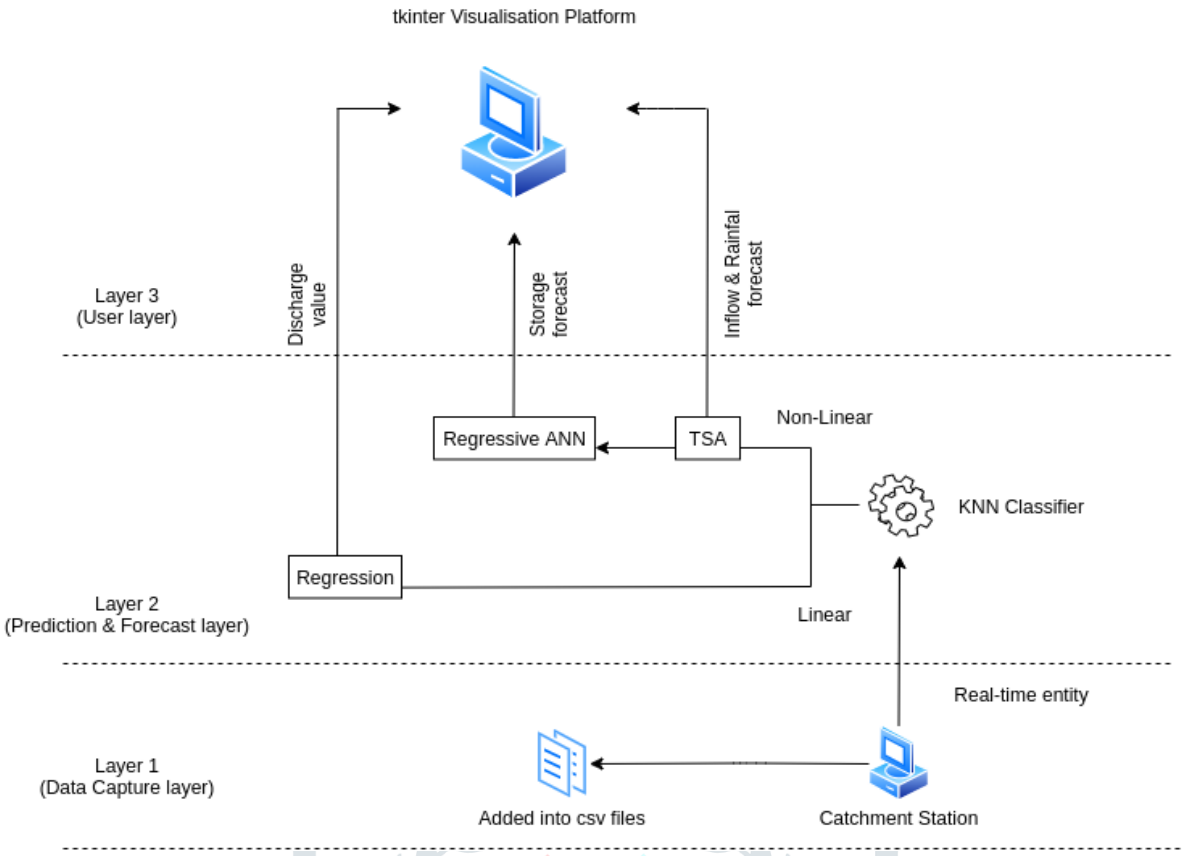

Fig 1. Architecture

1. **Data Capture Layer**: In this layer we collect data from various reservoir stations, this data is collected as real-time basis and is provided at per day interval. This data is then sent into next layers where it acts as input and one copy of the same date is recorded into csv files that will later be used to fine tune our models, this is our recursive learning feature.

2. **Prediction and Forecast Layer**:

   A. KNN Classifier:

   This is layer where the actual computations are done, at this layer the output from previous layer is first interfaced with a KNN classifier which classifies the input into one of the five classes that we defined in the training phases based upon their feature, the sole purpose of classifying the data into different classes was to ease the process analysis and designing separate models for different class datasets that are more classified and may perform better than a generalized solution for the entire dataset. At this layer the date is primarily classified into two separate classes which are linear and non-linear and depending upon this the further executions are carried out.

   B. Regression:

   If the data is classified as linear then it is sent into this layer for further execution. The linear class is nothing but situation where in the inflow and outflow form a linear curve, basically remaining proportional to each

other at any proportionality constant, as suggested by red plots in following graph.

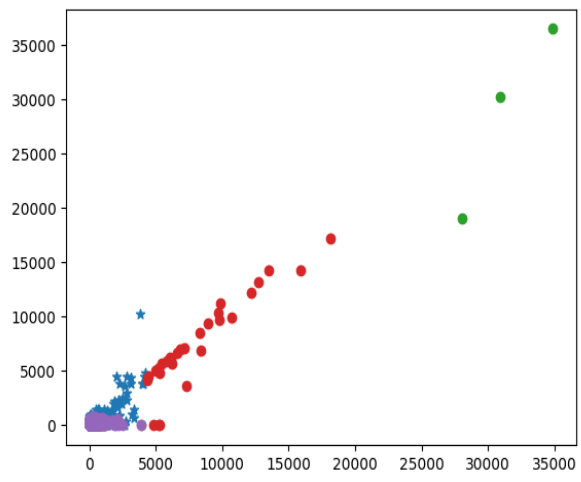

Fig 2. Feature space of outflow and inflow

This is when inflow and outflow have proportional rates and are unaffected by any other factor, specially rainfall. When there is no rainfall or any other impact of complex hydrological cycle then as the reservoir receives the inflow, it incorporates the discharge at same rate. This can be acknowledged by the fact that when inflow and discharge have a linear relation, the storage tends to remain constant and at a good level (illustrated in fig 3).

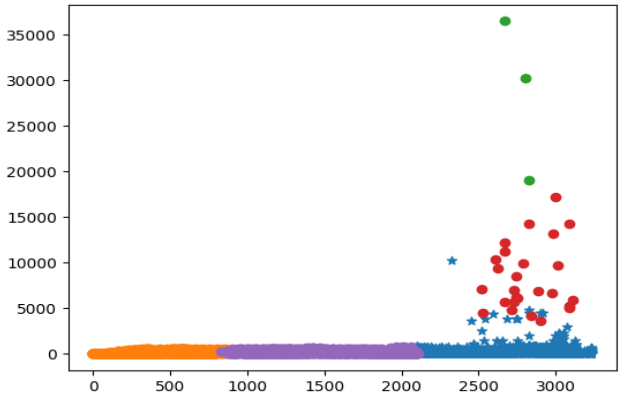

Fig 3. Feature space of outflow and storage

C. Time Series Analysis:

This is the module that takes on the control when output of classifier is any one of non-linear class i.e. when there is no linear relation between the two factors and the outflow has a variable dependence on other factors too. In this case prediction for an appropriate value of outflow becomes a bit more complicated process so we facilitate the decision process by providing some forecasted values of these parameters on a daily and a monthly basis. We applied time series analysis for two parameters, the recorded rainfall and river inflow. The real-time analysis of these factors was done by using recurrent neural network that took the real-time input from the station and used it to produce forecast for coming five days so that a prior analysis of certain vulnerabilities can be examined and tackled earlier.

Along with this we also trained an ARIMA based model that produced forecast for the same factors but these were trained from historical data and learned to map a function that could predict these values by considering seasonality and trends in the data collected over the time. This was also important for a reason that real-time analysis could never account for various complexities induced due hydrological cycles, one of which are monsoonal anomalies affecting the rainfall predictions and indirectly somewhere also affecting the inflow. So, to get a prior warning of monsoon arrival we needed to have model that could consider the seasonality in these datasets.

The output of RNN for rainfall and inflow made very good prediction with an MSE of around 144 – 160 for rainfall and around 1440 – 1600 for inflow. The following graphs does the comparison for predicted and actual recorded values of the parameters.

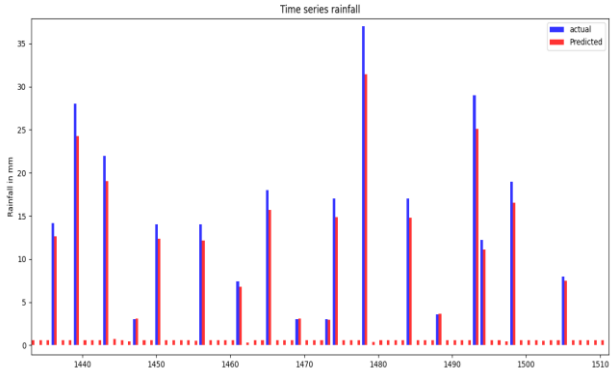

Fig 4. RNN time series analysis for Rainfall

The figure depicts the outcome of time series analysis using RNN for real-time rainfall prediction. The blue bar chart is the actual and red bar chart is the predicted value, as evident from the diagram they closely align with each other solidifying the accuracy of the model. The RNN time series maps both inflow and rainfall times series as a function $f(a)$: $V_i \rightarrow V_o$, where V_i: $[V_1, V_2, \ldots, V_n]$ and V_o: $[V_1, V_2, \ldots, V_m]$ here m and n are future time steps and time lag respectively. In our study we took values of m and n as 5, meaning we considered previous five time step values to present forecast for next five days.

This forecast was correct until it was in a particular hydrological cycle or in particular seasonal period, but during a transition period while moving from one season to other, these forecasts started showing a negative impact by making prediction with an accuracy lower than expected and eventually leading to a performance degradation. So, this required a building a solution that would not just make prediction based upon recent trends but also be able analyze historical trends and seasonality in these datasets. The most appropriate algorithm that could possibly defy the performance fallout at transition period was ARIMA for recording various seasonal trends to counter transitional effect by alarming about various shifts like monsoon arrival etc. And as our dataset was highly stationary it was most appropriate to use ARIMA at this level. This model functions in an exactly same way as RNN other than the fact that it considers larger time lags to analyze larger windows and learn about various seasonal factors.

Fig 5. ARIMA time series analysis for rainfall

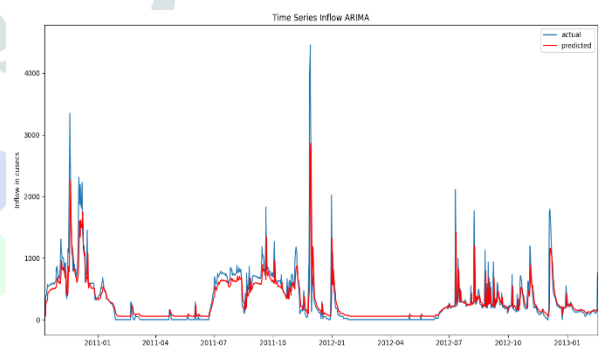

Fig 6. ARIMA time series analysis for inflow

And as per expectations the transition were clearly predicted as evident from above diagram (fig 5). So, the ARIMA was integrated along with the RNN as an alarming system for any transition and also as second perspective about these factors.

Along with a daily analysis of the rainfall there is one feature of monthly analysis of the rainfall added into the system just to make a more informed decision about water management processes. And also, as the size of the time window under consideration for time series analysis decreases the model tends to give a more accurate and faster results. So, for this purpose we had to make use of monthly rainfall data that had a higher seasonality and a lesser trend, therefore we

made use of Seasonal ARIMA (SARIMA) that functions very well with these considerations. As a monthly seasonal trend makes appearance after every 12 or 13 months, the SARIMA was trained accordingly on these parameters. In fact, the SARIMA model outstood any other daily analysis model in terms of accuracy. Hence it was integrated into the system, though it made forecasts for a comparatively bigger time window.

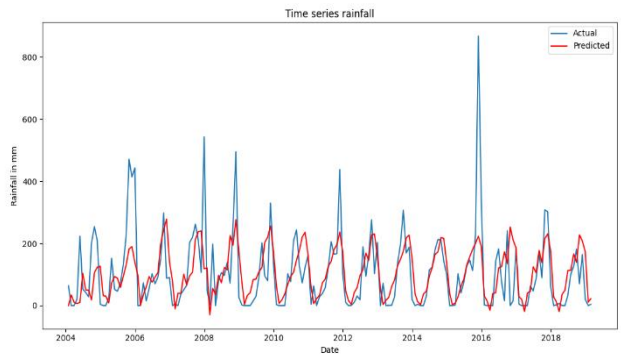

Fig 7. SARIMA for monthly rainfall

The figure depicts the actual and the fitted curves for monthly rainfall over the period of several years. The figure also explains the seasonality in form of spikes occurring in an interval of every 12 months.

D. Regressive Neural Network:

Till this point we didn't model anything for storage prediction of reservoir which is actually the most determining factor in this context and can prove to make the decision-making process more reliable. Unlike rainfall and inflow this wasn't meant to be predicted accurately using time series analysis or any of the previous models, it required a more sophisticated approach. As the datasets that were present with us were capable of mapping the parameters as it was recorded, it was never possible for the storage value to exceed its fixed capacity or the threshold even though it might have gone beyond it at some point definitely. It was not possible for the storage to be recorded in a tragic condition; hence it was never possible for a normal time series analysis model to predict any flooding condition in advance and similarly for any drying up situations. A similar model would have never mapped any function predicting storage causing overflow of the resources.

So, rather than using a regular time series analysis and to counter the challenges discussed in the previous section, we map storage level ANN based prediction process as a function, $f(a) : I_a \rightarrow O_a$ where,

$I_a : [S_i, I_i, R_i]$ is set representing input (S_i is the storage of the reservoir, I_i is the inflow, R_i is the rainfall recorded) and $O_a : [S_{i+1}]$ is the output of the model where S_{i+1} represents the storage of the reservoir a day ahead.

Basically, the model learns how current storage level is influenced by previous storage levels along with inflow and rainfall. We know the current level of storage is actually deviation in previous time steps storage value caused by increase due to inflow and rainfall and any differencing factor due to discharge that was released on that day. Other than the outflow, in the previous sections we have devised a model that can make a good estimate of inflow and rainfall values for future time steps which can in turn be used along with the recorded current storage levels, it is important to note here that as we do not have any estimate for outflow, the model being discussed here will have to learn any impact caused by outflow by itself as the model will be trained by the dataset that illustrates impact of current storage level, rainfall and inflow on future value of storage. The trained neural network therefore has an ability to predict future reservoir storage levels based on the current reading, recorded at the reservoir. With rigorous training of ANN, we were able to provide optimized results in shorter interval of time. The ANN prediction engine thus was able to provide a certain accuracy target.

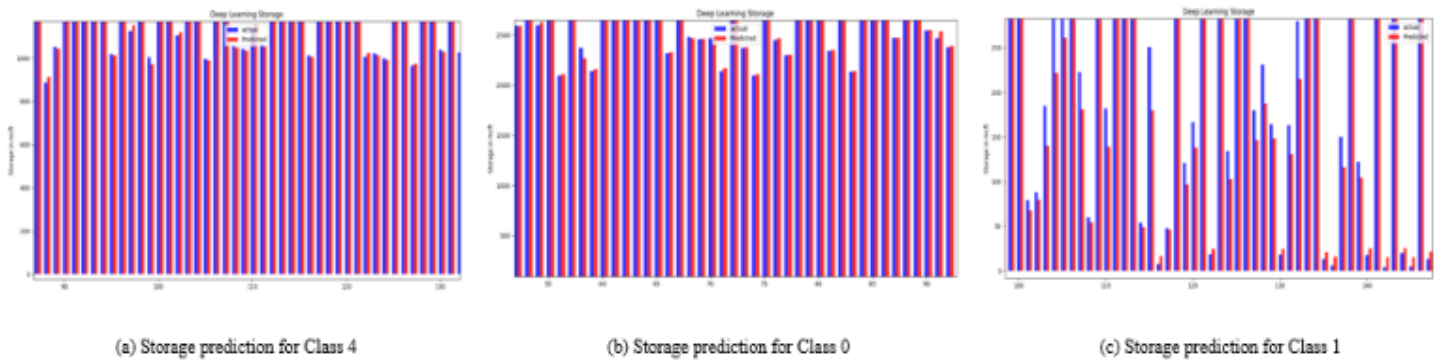

Fig 8. Outcome of Deep Neural Network for various classes for storage prediction

3. **User layer:** This layer represents a visual interface for users where the data from prediction layer are interpreted followed by some manipulation and then displayed to the user. The user layer represents a desktop platform where necessary analysis and information is presented to the user according to their profile. We enable user with the specific data of their interest depending upon their role in water management process. We use this desktop platform to assist user in process of decision making rather assist them with some highly relevant data to take more informed decisions. This is also the layer where we have integrated our mathematical support tool, we discussed earlier that would process the outcomes so far to approximate a value for discharge rate that aims to keep the storage at a neutral level. This mathematical model assumes inflow and rainfall to be directly proportional to outflow with storage being the proportionality constant, which basically we try to estimate from changes happening in these parameters. The relationship is formulated as,

$$(\text{inflow} + \text{rainfall}) \propto (\text{outflow})$$
$$\text{storage} = (\text{inflow} + \text{rainfall}) - (\text{outflow})$$

So, to maintain storage degradation rate it uses basic principle to adjust the rate of discharge over time in accordance to rate of inflow and rainfall and hence keeping storage at a constant level or at a precise rate of degradation. Later we also compare this discharge with demands of the consumers to refine the analysis. In this way user is provided with knowledge to observe the status of reservoir and make informed decision.

IV. RESULTS

Through our analysis we intend to predict some of the very relevant and important factors in water management processes by employing the methodology presented in previous sections and assisting the user to make informed decisions. We conducted various tests for different classes of data and recorded the outcome of the system and analyzed their accuracy and impact. The system outperformed in some of the attempts, but was able to produce desired accuracy in most of the cases. We used our to generate results for 2018's October month and compared generated outcome with actual values of these parameters. The following diagram represents the output generated by the user layer of the system, as there were no potential alarming conditions like, shift into new hydrological cycle there were no warning displayed into the system.

Identified class label : 1

Day Wise Analysis : Real-time

Parameter	[Initial Data]	[Day 1 prediction]	[Day 2 prediction]	[Day 3 prediction]	[Day 4 prediction]	[Day 5 prediction]
Rainfall (mm)	27	13.93	23.08	12.39	19.99	11.05
Inflow (cusecs)	801	626.13	804.46	633.36	808.06	639.74
Storage (mcft)	317	319.44	286.28	324.48	344.05	388.22

Day Wise Analysis : Seasonal

Parameter	[Initial Data]	[Day 1 prediction]	[Day 2 prediction]	[Day 3 prediction]	[Day 4 prediction]	[Day 5 prediction]
Rainfall (mm)	27	38.69	27.97	20.87	17.81	11.84
Inflow (cusecs)	801	640.81	579.98	573.57	581.95	619.75
Storage (mcft)	317	319.44	373.44	324.11	404.8	442.12

Approximate Discharge Value (cusecs): 37.05

Fig 9. Output phase of system for class 1 dataset.

Monthly Analysis : Rainfall

	[current total]	[predicted total]	[next month]
Value (mm)	0.0	161.05	169.46
Curve	view	view	view

Fig 10. Output phase for monthly analysis feature for rainfall

V. CONCLUSION

Our paper presents an intelligent system for water optimization process that uses, deep neural network based prediction engine for various crucial factor in water operations. Our solution provides useful insights into the domain of water optimization and hence assisting the user to take more informed decisions and preserve the hydrological ecosystems. The system makes collective use of real-time data and historical data to not only make prediction about certain factors but also amounts the most suitable value for discharge rates by considering the pattern in which water is consumed by various habitants thus lowering the risk of degradation and malfunctioning of these water bodies.

We have focused our study to Poondi Reservoir, Tiruvallur, Tamil Nadu which is one of the main sources of supply to the Chennai metropolis. But the solution can be generalized for any reservoir by performing same kind of analysis and devising models that maps to trends and pattern in particular reservoirs data. By this approach we also remove the use of slower and less reliable mechanical processes build upon IoT with high performance computational systems making use of deep neural networks.

REFERENCES

[1] Ioan Petri, Member, IEEE, Baris Yuce, Member, IEEE, Alan Kwan, Member, IEEE and Yacine Rezgui, Member, IEEE "An intelligent analytics system for real-time catchment regulation and water management" August 2018

[2] Ioan Petri, Javier Diaz-Montes, Mengsong Zou, Ali Reza Zamani, Tom Beach, Distributed Multi-Cloud Based Building Data Analytics, Kecskemeti, Gabor, Attila Kertesz, and Zsolt Nemeth." Developing Interoperable and Federated Cloud Architecture." IGI Global, 2016. 1-398. Web. 24 Feb. 2017. doi:10.4018/978-1-5225-0153-4

[3] Baris Yuce, Haijiang Li, Yacine Rezgui, Ioan Petri, Bejay Jayan, Chunfeng Yang, utilizing artificial neural network to predict energy consumption and thermal comfort level: An indoor swimming pool case study, Energy and Buildings, Volume 80, September 2014, Pages 45-56, ISSN 0378-7788.

[4] N. Chattha, R. Karki, Senior Member, IEEE, F. Fang Power Systems Research Group University of Saskatchewan Saskatoon, Canada "Seasonal Reservoir Management in Hydro Dominant Power Systems to Enhance Availability" 2017

[5] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, Internet of Things (IoT): A vision, architectural elements, and future directions, Future Gener. Comput. Syst. vol. 29, no. 7, pp. 1645-1660, Sept. 2013.

[6] T. Robles et al.," An Internet of Things-Based Model for Smart Water Management," Advanced Information Networking and Applications Workshops (WAINA), 2014 28th International Conference on, Victoria, BC, 2014, pp. 821-826, doi: 10.1109/WAINA.2014.129

[7] G.-H. Wei, F. Liu, and L. Ma, "Fuzzy optimization of water resources project scheme based on improved grey relation analysis," in Proc. 3rd Int. Conf. Comput. Res. Develop., Shanghai, China, 2011, vol. 4, pp. 333–336.

[8] B. Huang, F. Gui, and X. Zhang, "Study of the optimal water resources allocation scenarios in Pingxiang city," Adv. Intell. Soft Comput., vol. 105, pp. 487–493, 2011

[9] Singh et al, Appropriate data normalization range for daily river flow forecasting using an artificial neural networks, Hydroinformatics in Hydrology, Hydrogeology and Water Resources. International Association of Hydrological Sciences Publication, Uk, pp. 51-57, 2009

[10] W. C. Wang, K. W. Chau, C. T. Cheng, L. Qiu, A Comparison of Performance of Several Artificial Intelligence Methods for Forecasting Monthly Discharge Time Series, Journal of Hydrology 374(3-4) (2009) 294-306.