Reservoir release forecasting by artificial neural network at Pa Sak Jolasid Dam

Thammanoon Thepsena
Department of Computer Engineering, College of innovative technology and engineering
Dhurakij Pundit University
Bangkok, Thailand
647191070006@dpu.ac.th

Chaiyaporn Khemapatapan
Department of Computer Engineering, College of innovative technology and engineering
Dhurakij Pundit University
Bangkok, Thailand
chaiyaporn@dpu.ac.th

Dendej Rakrattanatrai
Department of Computer Engineering, College of innovative technology and engineering
Dhurakij Pundit University
Bangkok, Thailand
dendej.rak@dpu.ac.th

Abstract— **Due to floods in the past year The Pa Sak River Basin is therefore another watershed source that causes widespread damage to life and property. An interesting predictive study of the dam release is therefore an experimental idea. Using data recorded at the beginning of '64 using artificial neural networks techniques, compared with a support vector regression model. Estimated for satisfactory results by an ANN precision model R^2 approximately 0.88513 RMSE approximately 8.044 and the SVR model R^2 result. Approx. 0.43369 RMSE Approx. 18.484.**

Keywords- Reservoir discharges; Pa Sak River; Pa Sak Jolasid Dam; Discharges forecasting; Artificial neural network

I. INTRODUCTION

At present, it is a wild water flowing from the mountains in Phetchabun, Loburi and Chaiyaphum provinces from the influence of Typhoon Dien Mu to the Pa Sak river. As a result, the amount of water in the Pa Sak dam increased rapidly. Within a week, water volumes increased from 44% of the basin's capacity on Sept. 22 to nearly 107% of the total water volume. Basin capacity today (1 Oct., 2021) statistics from the Royal Irrigation Department also indicate that the amount of water in the Pa Sak Reservoir as of October 1 this year was up to 1,027 million cubic meters. The highest compared to the same days of 2018, 2019 and 2020, which were measured, the water volume in the reservoir was 680 million, 439 million and 375.8 million cubic meters, respectively. The dam is 4,860-meter (15,940 ft) wide and 36.5-meter (120 ft) high dam is earth-filled with an impervious core. Fig.1 the storage capacity is 785 million m^3 of water at normal water level, with a maximum capacity of 960 million m^3. The dam also supplies about 6.7 MW of hydro-electric power.

The main purpose of the downstream drainage is to allow the Pa Sak Reservoir to support the amount of water above which continues to flow in continuously. and to maintain the stability, safety, and security of the dam National Water Resources Office as deputy director of the National Water administration said that at present there are still about 400 million cubic meters that will flow into the Pa Sak Reservoir.

More but because between 28 Sept.,2021 - 2 Oct.,2021 the monsoon trenches will move down across the lower central and southern regions. As a result, there was heavy rain in some areas. Therefore, there may be water flowing into the basin continuously.

The Royal Irrigation Department explains that the amount and speed of water from the reservoir into the Pa Sak River will be set at a level that does not cause too much impact on the downstream area or is at a level that "The Pa Sak river downstream can support water. Without overflowing the embankment of the Pa Sak river" from the Pa Sak Dam in Phatthana Nikhom District, Lopburi province, the water mass drained from the dam will flow into the Pa Sak river through Wang Muang District, Kaeng Khoi District, Mueang Saraburi District, Saraburi District. Sao Hai of Saraburi province, then it will reach Rama 6 Dam in Tha Ruea District, Phra Nakhon Si Ayutthaya province.

The water management of the Pa Sak Dam in emergency and urgent situations in order to minimize the impact on all parties must be done in this kind of situation. This study used data on water inflow rates. the water level in the basin Reservoir storage volume and rainfall. Such data were collected from January 1, 2021, to November 2, 2021, by the Reservoir Water Database System, the Royal Irrigation Department to help analyze the probability of water release using Artificial intelligence technology studies behavior and train models to learn from the recorded. To accurately predict tides play an important role in the effective management of water resources during droughts and floods. water storage and reservoir operations traditional methods have been studied and models for predicting tides and rainfall can be divided into two categories: physical methods and empirical methods [1][2]. physical form forecasts are generated by models built from system equations to predict tides or rainfall. Related weather forecasting systems such as temperature changes precipitation forecast and changes in water pressure are used as mathematical equation variables to create forecasting models. The application of these models depends on the region studied and the data available for that region. Physical modeling for better illustration.

2022 7th International STEM Education Conference (iSTEM-Ed) | 978-1-6654-9821-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISTEM-ED55321.2022.9920896

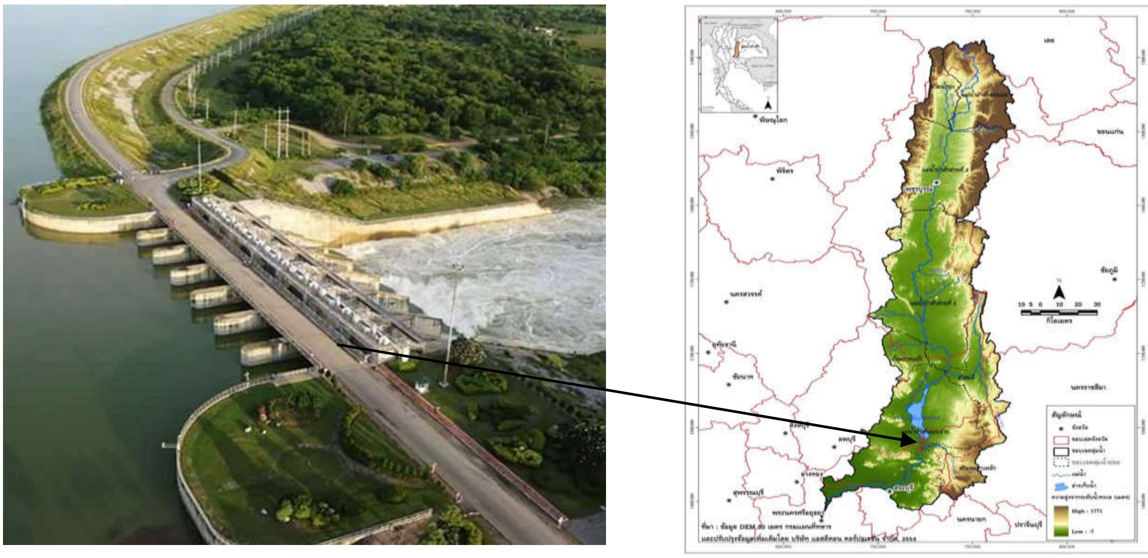

Fig.1 Pa Sak Jolasid Dam and watershed area 15,625.87 Km^2

It provides comprehensive and straightforward information about the processes involved in predicting tides. These models are used as parameters that are directly linked to watersheds. However, the predictability of physical models requires comprehensive data. which sometimes may be difficult to find in the study area [3]. The empirical approach is based on an analysis of historical aquatic climate data and its relationship to a much wider range of areas. Empirical techniques involve statistical models or classical machine learning models such as linear regression. Support vector machine fuzzy logic Nearest neighbor average, etc. [4]

Machine learning models, part of the widely used artificial intelligence technology, were used in this study. and then used to calculate the accuracy of the model. to use as a hypothesis and reference in further in-depth study this study used the dataset in the Fig.2 as a relationship. before making a forecast data are inserted into the model multilayer perceptron neural network according to table1, a dataset the feature to rain, storage, w inflow and set the target to release, let the model train, test the data from a feature, target. In this study, train 70%. and test 30% of the items recorded at the beginning of 2021 of 294 rows × 4 columns.

The remainder of this document follows where we describe the relevant work in Part 2. while our dataset is covered in Part 3. The models used in this work are reported in Part 4. The results are discussed in Part 5 and conclusions in Part 6.

II. RELATED WORK

In the study of machine learning algorithms and especially neural networks. (ANN) for use in applied hydrology It has been studied for a long time such as Dahamsheh and team [5] have successfully implemented rainfall forecasting. Callegari's team[8] predicted river flows using the Support Vector Regression (SVR) technique in 14 reservoirs as monthly data in northern Italy. Effective results are obtained compared to long-term and linear models. They show that the SVR model has RMSE (Root Mean Square Error) = 22% on a 1-month forecast. [9, 10], Centuries of meteorological hydrological analysis for the Adda Basin (Central Alps). The basis of the nervous system ANNs are composed of multiple layers of neurons. It consists of an input layer, a hidden layer which can contain more than one layer, and an output layer. For creating high-performance models to predict outcomes in supervised regression, the input data is directed into the system for training, validation, and testing procedures. Used to calculate and upgrade network elements such as weights and bias. Validation procedures or evaluation of the effectiveness of the training process. After that, it will be the final test of the system.

	Feature			Target
Date_stamp	**rain**	**storage**	**inflow**	**release**
0	0	653.7	0	3.03
1	0	639.9	0	1.3
2	0	637.1	0	2.17
3	0	634.3	0	2.17
4	0	631.6	0	3.47
...	...	...	...	...
289	290.5	980	35.84	34.59
290	226.5	982	34.64	30.31
291	1	989.8	38.52	30.31
292	1	998	38.53	30.32
293	64	1011	43.43	30.3

Table 1 Data frame for prediction model

Feed Forward neural networks

The forward arrangement and connection produce static ANNs, the simplest type of neural network structure. The input signal travels to the input layer with weights and bias, passes through the hidden layer, and also contains weights and bias in it. before being passed to the output layer. Without any loops, Feed Forward Neural Networks (FFNN) is a well-known construct. It is practical and has the simplest structure as in Equation (1). A two-layer feedforward neural network with n input units, m output units and N units in the hidden layer, is shown in Fig. 2

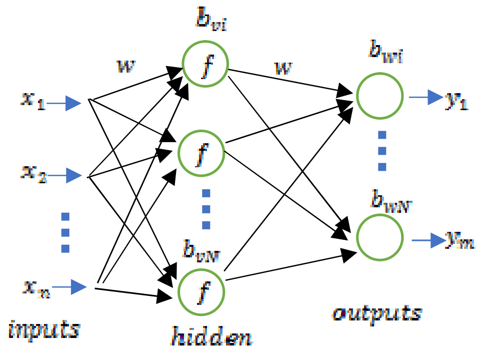

Fig.2 Feedforword Neural Network

The output vector y is determined in terms of the input vector x by the formula

$$y_i = \sum_{j=1}^{N} [w_{ij} \cdot f\left(\sum_{k=1}^{n} v_{jk} x_k + b_{vj}\right) + b_{wi}] ; i = 1, \ldots, m \quad (1)$$

where f are the activation functions of the neurons of the hidden layer. The inputs to hidden layer intercom-nection weights are denoted by v_{jk} and the hidden layer to outputs interconnection weights by w_{ij}. The bias weights are denoted by b_{vi}.

Support Vector Regression

Support Vector Regression (SVR) is a variant of the support vector machine (SVM) algorithm that distinguishes data belonging to different classes. using hyperplane line This hyperplane is the ultimate line between data classes. If the data cannot be linearly split, the SVM uses a non-linear function called kernels, to automatically translate instances of the training data, Support Vector Regression (SVR) into a type of SVM applied to real numbers. In the case of regression The margin of tolerance or maximum error is set to suit the desired accuracy. The idea is to minimize errors. by the hyperplane, which must calculate the margin bearing in mind that some errors are acceptable. Michelle Pini and colleagues [11] conducted an experiment. Reservoir forecast in Italy SVR is one of the experiments. Compare to Other models The experiment yielded quite accurate results.

III. Data set preparation

From the foregoing above The data from this experiment, recorded from National Water Resources Office in January 1, 2021, to November 2, 2021, are shown in Table one to three column are Features and one Target with 294 rows.
Interesting features are

- Rain fall
- Storage of dam
- inflow to dam

The water release data during that time period is used to forecast the model. when bringing Feature In which we are interested in the mapport graph compared to the recording time period, the time-of-year relationship to water management is shown in Figures 3. show water management in the conditions noted above as a relationship. of the amount of water inflow into the dam, the amount of water retention, and the depletion of the dam under normal and high-water situations.

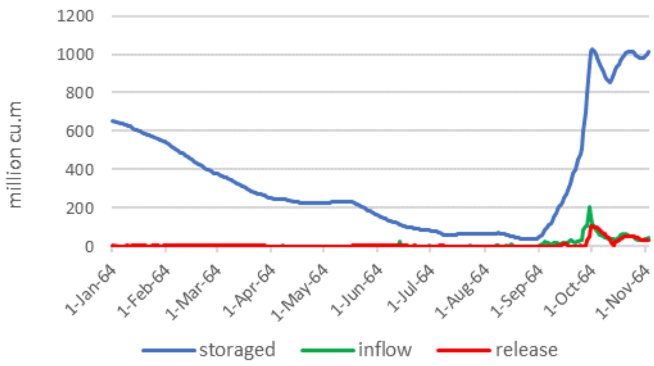

Fig. 3 Actually water management

Figures 4. Show rainfall throughout the period of measurement, it can be seen that there is a lot of rain during the storm in late October. and related to the amount of water flowing into the Pa Sak Dam.

Fig. 4 Amount of rainfall

In the experimental process, Feature and Target data were examined for accuracy, Clean missing data and making the training/test split with approximately a 70/30 ratio. these models is assessed using two metrics namely the root mean squared error *(RMSE)* and the coefficient of determination *(R^2)* estimated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \qquad (2)$$

$$R^2 = 1 - (\sum_{i=1}^{n}(\hat{y}_i - y_i)^2)/(\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2) \qquad (3)$$

IV. MODEL USED.

An experiment using the principles of ANN and SVR will be done in the form shown in Figure 5. is to take the collected data, select columns of data set, clean missing data, then extract the data to get the data to train 2 parts of 70 percent, the other 30 percent will be tested. Both Feature and Target, when the model are processed, takes this information for accuracy and discrepancy. This will show the difference when comparing ANN and SVR.

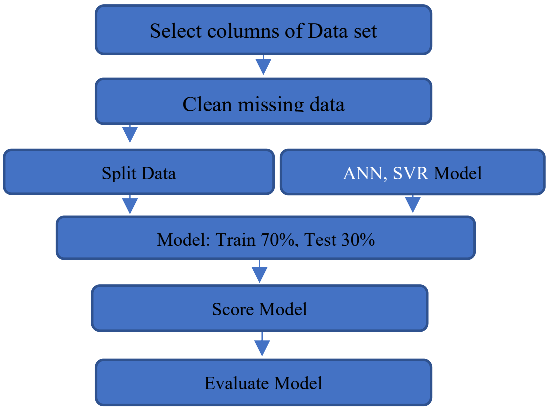

Fig. 5 Overview of research methodology of the study

V. RESULTS AND DISCUSSION.

The result of predicted and actual for ANN model is shown in Fig. 6 scatter plots on test set in x-axis actual values, in y-axis predicted values. In blue line of perfect agreement. The blue dot shows the expected number of samples/instances predicted and aims to find predictions near the diagonal blue line. However, an analysis of the results by ANN found that the projections were closer to the diagonal blue line than the SVR model in Fig. 7 and, considering the ANN's R2 accuracy, it was 0.88513. and a tolerance of RMSE 8.044, while the SVR model gave R2 0.43369 and RMSE 18.484.

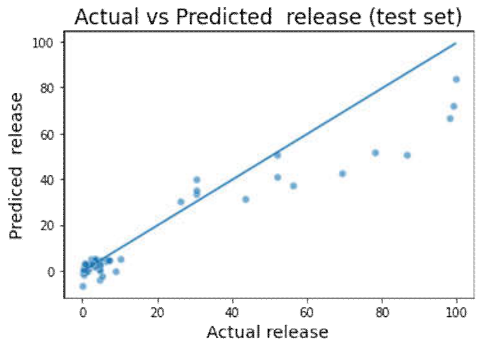

Fig. 6 the result of predicted and actual ANN model

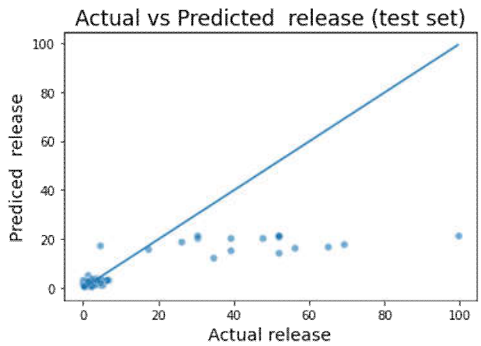

Fig. 7 the result of predicted and actual SVR model

VI. CONCLUSION.

Reliable estimation of hydrological parameters is essential for managing and preventing natural hazards. Tides forecasting plays an important role in regulating the amount of water. Data-driven models provide insights that help make good decisions. Hydrological data show a nonlinear data model. in this study Various machine learning models were examined. for effective forecasting of currents in dams We looked at different statistical measures to assess the model's performance. We found that ANN showed outstanding results compared to SVR. This is due to ANN's ability to learn nonlinear data models. In the future, we will consider other meteorological variables such as (farmers' water needs, La Niña and El Niño phenomena, wind speed and solar radiation), which are important variables affecting all currents. in Pa Sak Jolasid Dam

REFERENCES

[1] C. Refsgaard, J. Knudsen, Operational validation and intercomparison of different types of hydrological models, Water Resources Research32 (7) (1996) 2189–2202.

[2] M. Valipour, M. E. Banihabib, S. M. R. Behbahani, Comparison of the arma, arima, and the autoregressive artificial neural network models in forecasting the monthly inflow of dez dam reservoir, Journal of hydrology 476 (2013) 433–441.

[3] H. Karimi, H. Zeinivand, N. Tahmasebipour, A. Haghizadeh, M. Miryaghoubzadeh, Comparison of srm and wetspa models efficiency for snowmelt runoff simulation, Environmental earth sciences 75 (8) (2016) 664.

[4] G. Uysal, A. U. S¸orman, Monthly streamflow estimation using wavelet-artificial neural network model: A case study on c¸amlıdere dam basin, ¨turkey, Procedia computer science 120 (2017) 237–244.

[5] A. Dahamsheh, H. Aksoy, Markov chain-incorporated artificial neural network models for forecasting monthly precipitation in arid regions,Arabian Journal for Science and Engineering 39 (4) (2014) 2513–2524.

[6] E. Toth, Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting, Hydrology and Earth System Sciences 13 (9) (2009) 1555.

[7] W. Chen, X. Xie, J. Wang, B. Pradhan, H. Hong, D. T. Bui, Z. Duan, J. Ma, A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility, Catena 151 (2017) 147–160

[8] M. Callegari, P. Mazzoli, L. De Gregorio, C. Notarnicola, L. Pasolli, M. Petitta, A. Pistocchi, Seasonal river discharge forecasting using support vector regression: a case study in the italian alps, Water 7 (5) (2015) 2494–2515.

[9] A. Crespi, M. Brunetti, R. Ranzi, M. Tomirotti, M. Maugeri, A multi-century meteo-hydrological analysis for the adda river basin (centralalps). part i: Gridded monthly precipitation (1800–2016) records, International Journal of Climatology (2020).

[10] R. Ranzi, E. Michailidi, M. Tomirotti, A. Crespi, M. Brunetti, M. Maugeri, A multi-century meteo-hydrological analysis for the adda river basin (central alps). part ii: daily runoff (1845–2016) at different scales, International Journal of Climatology.

[11] Michele Pini, Andrea Scalvini, Muhammad Usman Liaqat, Roberto Ranzi, Ivan Serina, Tahir Mehmood, "Evaluation of Machine Learning Techniques for Inflow Prediction in Lake Como, Italy," Procedia Computer Science, Volume 176, 2020, pp. 918-927, ISSN 1877-0509,