Mathematical Modelling for Short-Term Urban Water DemandForecasting Using Wavelet Support Vector Regression

Amrit Kaur[1], Jagroop Kaur[2]

[1,2]Department of Computer Science &Engineering, Punjabi University, Patiala, India
Email:amrit_ce@pbi.ac.in[1], jagroop_ce@pbi.ac.in[2]

Received: 15.10.2023 Revised: 28.11.2023 Accepted: 21.12.2023

ABSTRACT
This study presents an approach for short-term urban water demand (UWD) forecasting that couple's wavelet transform with Support Vector Regression (SVR). The wavelet transform decomposes the original urban water demand time series into multiple sub-series at various frequency levels, capturing long-term trends and short-term fluctuations. These decomposed sub-series serve as inputs for the SVR model, identifying complex nonlinear relationships among diverse factors influencing water demand. The model's performance was evaluated forlead periods of 1 and 3 days, with results demonstrating that the Wavelet-SVR model outperforms the standard SVR method in forecasting urban water demand. Performance metrics, including Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), were used to quantify the accuracy of the predictions.

Keywords: Urban water demand, Support vector regression, Wavelet transform

INTRODUCTION
Predicting water demand is essential for efficiently designing and managing water supply systems. It facilitates activities like planning for future expansions or system improvements, optimizing the capacity and operation of reservoirs, pumping facilities, distribution networks and ensuring the effective management of urban water resources. Short-term forecasting deals with small-scale planning and management, typically covering timeframes from one hour to afew days. This approach includes hourly or daily predictions, helping utilities to plan for and address water demands for upcoming events. [1]. In contrast, long-term forecasts such as monthly or yearly estimates are essential for strategic planning and infrastructure development [2]. Predicting consumption patterns allows utilities to minimize wastage, enhance energy efficiency, and respond more effectively to unexpected events like droughts or peak usage. Various mathematical models, such as regression analysis, time series forecasting, and machine learning methods, have been created and applied in multiple regions worldwide to forecast urban water consumption. The study of urban water demand forecasting incorporates various techniques, including linear statistical approaches and non-linear models, suggested by[3], [4],who examines the extensive array of models presented in the research while highlighting their strengths and limitations. Linear regression models are widely acknowledged as a foundational method in predictive analysis, valued for their simplicity and interpretability. However, they often need to catch up in accurately capturing the complex nonlinear relationships in water demand data [5], [6].
Machine learning (ML) forecasting models are gaining recognition in hydrology due to their ability to adapt to non-linear data dynamics [7], [8]. In various case studies, artificial neural networks (ANNs) have consistently surpassed conventional forecasting techniques, including multiple linear regression (MLR) and time series analysis. In research conducted in Ottawa, Canada, ANN models proved superior accuracy in predicting peak daily summer water demand compared to MLR and time series methods [8]. Subsequent investigations conducted in Kanpur, India, show that ANN models consistently outperformed regression and time series approaches, achieving an average forecasting error of only 3.28% [5].
The support vector regression (SVR) variant of the support vector machine represents a viable machine-learning technique for predicting water demand. By leveraging kernel functions, SVMs often achieve superior performance compared to artificial neural networks (ANNs) in capturing non-linear data patterns, resulting in enhanced accuracy in forecasts [9], [10]. A study reported that a Support Vector Machine (SVM) model optimized using a genetic algorithm achieved greater forecasting accuracy than

ANN in Ziyang City. This finding indicates that, with proper optimization, Support Vector Regression (SVR) can outperform ANN in short- term forecasting scenarios. Based on the normalized mean square error (NMSE) criteria, SVR models have proven more effective than traditional back propagation neural networks in predicting drought [11].

When not adequately addressed through proper data preprocessing, seasonal patterns can substantially degrade the performance of AI models in accurately capturing and managing non- stationary dynamics. Hybrid models are increasingly utilised in water demand forecasting because they can overcome standalone models' limitations, offering enhanced accuracy and reliability. [12] suggested integrating multiple techniques to exploit their complementary strengths while mitigating individual weaknesses. Hybrid structure wavelet analysis with SVR and particle swarm optimization (WA-PSO-SVR) coupling outperformed standalone SVR models in predicting water quality indicators, particularly for extreme values and short-term forecasts [13].

The study aims to forecast urban water demand (UWD) for one and three-day lead times by employing a Discrete Wavelet Transform (DWT) for time-series decomposition and extracting critical features from different frequency components. A hybrid forecasting model is then developed by integrating wavelet-based preprocessing with Support Vector Regression (SVR). The performance of the proposed model is evaluated using metrics Mean Absolute Error (MAE), Root Mean Squared Error (RMSE)

Research Material and Data Collection

In this study, UWD forecasting was performed for the Model Town region of Patiala, Punjab, India, with a population of 5,925. The water distribution system in this area operates with an average discharge of 0.8 million litres per day (ML/d), and a supply duration of 0.35 days. Thedaily bulk water demand is 0.4 million litres (ML), while the storage tank has a capacity of 1.2 ML. The recorded demand, with a 1-day lead time, was 0.76 ML, and the average daily per capita water consumption was approximately 135 litres.

The analysis was based on water consumption data collected between June 2018 and June 2023, sourced from water bills issued during this period. Key input variables used in the forecasting models included Urban Water Demand (UWD), air temperature (TEMP), humidity (HMD), wind speed (WS), air pressure (AP), weekend indicators (Wknd), holiday (HD) indicators (a binary variable denoting holidays), and population (P). This dataset was leveraged to develop accurate models for UWD prediction, intending to optimize short-term water supply management and address demand fluctuations effectively.

RESEARCH METHODOLOGY

The research methodology focuses on implementing the WSVR model for short-term urban water demand forecasting. The WSVR algorithm illustrated in Figure 1 was applied after preprocessing the dataset. Wavelet decomposition was employed to capture key patterns and reduce noise in the time series data, enhancing its representation of complex, non-stationary trends. The WSVR model was trained using the decomposed data to predict urban water demand over multiple lead times. Model performance was assessed using metrics such as Mean Absolute Error (MAE), and the results were compared with the standard SVR model to demonstrate the improvements achieved through wavelet enhancement.

1: Dataset

- X = {TEMP, HMD, WS, AP, Wkd, Wknd, HD, P} be the feature set.
- y=UWD be the urban water demand at time t
- t denotes hourly time index.
- The dataset $\{(X_t, y_t)\}_{t=1}^{T}$ consists of time indexed feature target pair.

2: Decompose the time series into using Daubechies Wavelet

- For each feature $X_i(t)$, decompose using the discrete wavelet transform (DWT)
  $X_i(t) = A_i(t) + D_t(t)$ (uptp level 1)
  where
  $A_i(t)$ is the low frequency (approximation) component of the feature set
  $D_t(t)$ is the high frequency (detail) component of the feature set
  i ∈ {TEMP, HMD, WS, AP, Wkd, Wknd, HD, P}
- Feature set after wavelet decomposition
  $X_t^{wavelet} = \{ A_{TEMP}, D_{TEMP}, A_{HMD}, D_{HMD} \ldots\ldots A_P, D_P$

3: Support Vector Regression (SVR) is applied to the decomposed features

- SVR fits a function f(X) that minimizes the error
- $f(X) = W^T\Phi(X) + b$
  where
  ϕ (X) is a kernel function mapping the input features into a higher-dimensional space.
  W are the model weights.
  b is the bias term.
- Minimize the regularized empirical risk
  $\min (W, b)[\frac{1}{2}||W||^2 + C\sum_{t=1}(\xi_i + \xi_t^*)]$
- Subject to constraints
  $y_t - W^T\Phi(X_t^{wavelet}) - b \leq \epsilon + \xi_t$
  $W^T\Phi(X_t^{wavelet}) + b - y_t \leq \epsilon + \xi_t^*$
  $\xi_t, \xi_t^* \geq 0$
  where
  C is the regularization parameter
  ϵ is the margin of tolerance
  ξ_i and ξ_t^* are slack variables that allow deviations from the margin

4: The Radial Basis Function (RBF) kernel is used for SVR

$K(X_t, X_j) = \exp(-\gamma||X_i - X_j||^2$
Where
γ is the kernel parameter

5: After training SVR model the forecasted water demand **for the next time stamp t+1 is:**

$\hat{y}_{t+1} = f(X_{t+1}^{wavelet}) = W^T\Phi(X_{t+1}^{wavelet}) + b$

6: The performance of the model is evaluated using MAE and RMSE

Figure 1: Framework of WSVR Algorithm

RESULTS AND DISCUSSION

In this study, SVR and WBSVR models were applied for forecasting UWD in Patiala, a city innorthern India. These models aimed to predict UWD with lead times of 1 and 3 days. Figure 2a and 2b Illustrate the performance of SVR and WSVR for both 1-day and 3-day ahead horizons, comparing the actual demand with the model forecast. Figure 2(a), representing the1-day ahead predictions, the WSVR model's forecast closely follows the actual demand, whilethe SVR model tends to overestimate the demand. This pattern holds in Figure 2(b), representing the 3-day ahead predictions, where WSVR consistently outperforms SVR in tracking the actual demand.

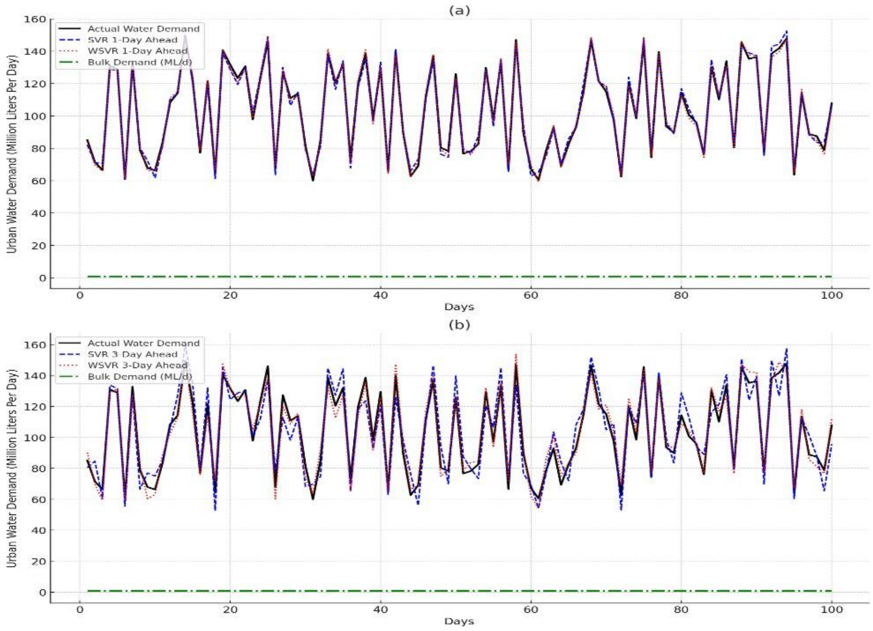

Figure 2: Actual and predicted water demand using SVR and WSVR (a) 1-day lead period b) 3-day lead period

Table 1 compares MAE and RMSE, further highlighting the accuracy of the WSVR model. For one day ahead, WSVR achieved a lower RMSE of 0.038 compared to SVR's 0.047 ML, and for three days ahead, WSVR's RMSE was 0.076, again lower than SVR's 0.092. The WSVR model shows a lower MAE value than the SVR model, indicating better performance.These metrics confirm that WSVR provides a more reliable forecast, particularly shorter-termpredictions.

Table 1: Model Performance of SVR and WSVR for the Research Region

Model	**Day 1**		**Day 3**	
	RMSE	**MAE**	**RMSE**	**MAE**
SVR	0.047	0.1990	0.092	0.1620
WSVR	0.038	0.1624	0.076	0.1308

CONCLUSION

The WSVR model demonstrates a superior ability to forecast urban water demand compared to SVR. Its closer alignment with actual demand values, both in the 1-day and 3-day ahead horizons, suggests that WSVR is better suited for accurate short-term urban water demand forecasting. This increased precision is critical for urban water management, as it allows for more efficient resource allocation and infrastructure planning, avoiding the inefficiencies caused by overestimated demand. Future work could further enhance the models by incorporating external variables like seasonal changes or population growth to refine predictions.

REFERENCES

[1] W. Hao, A. Cominola, and A. Castelletti, "Comparing Predictive Machine Learning Models for Short- and Long-Term Urban Water Demand Forecasting in Milan, Italy," IFAC-PapersOnLine, vol. 55, no. 33, pp. 92–98, 2022, doi: 10.1016/j.ifacol.2022.11.015.

[2] Ghiassi; M., D. K. Zimbra;, and H. Saidane, "Ghiassi2008.Pdf," vol. 134, no. 2, pp. 138–146, 2008, doi: 10.1061/_ASCE_0733-9496_2008_134:2_138.

[3] R. J. A. Donkor Emmanuel A, Mazzuchi Thomas A, Soyer Refik, "Urban Water Demand Forecasting: A Review of Methods and Models," Journal of Water Resources Planning and Management, 2012, doi: 10.1061/(ASCE)WR.1943-5452.0000314.

[4] S. Ghannam and F. Hussain, "Short-term water demand forecasting: a review," Australian Journal of Water Resources, vol. 00, no. 00, pp. 1–19, 2024, doi: 10.1080/13241583.2024.2350102.

[5] A. Jain, A. K. Varshney, and U. C. Joshi, "Short-term water demand forecast modelling at IIT Kanpur using artificial neural networks," Water Resources Management, vol. 15, no. 5, pp. 299–321, 2001, doi:

10.1023/A:1014415503476.
[6] M. S. B. A. Das and G. P. Pradhan, "A multivariate econometric approach for domestic water demand modeling : An application to Kathmandu , Nepal," pp. 573–589, 2007, doi: 10.1007/s11269-006-9030-6.
[7] J. Bougadis, K. Adamowski, and R. Diduch, "Short-term municipal water demand forecasting,"
[8] Hydrological Processes, vol. 19, no. 1, pp. 137–148, 2005, doi: 10.1002/hyp.5763.
[9] J. F. Adamowski, "Peak daily water demand forecast modeling using artificial neural networks,"Journal of Water Resources Planning and Management, vol. 134, no. 2, pp. 119–128, 2008, doi: 10.1061/(ASCE)0733-9496(2008)134:2(119).
[10] I. S. Msiza, F. V. Nelwamondo, and T. Marwala, "Water demand prediction using artificial neural networks and support vector regression," Journal of Computers, vol. 3, no. 11, pp. 1–8, 2008, doi: 10.4304/jcp.3.11.1-8.
[11] M. Herrera, L. Torgo, J. Izquierdo, and R. Pérez-García, "Predictive models for forecasting hourly urban water demand," Journal of Hydrology, vol. 387, no. 1–2, pp. 141–150, 2010, doi: 10.1016/j.jhydrol.2010.04.005.
[12] G. Wen, J. Yin, X. Zhang, and Y. Jin, "Application research of support vector machines in dynamical system state forecasting," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 5226 LNCS, pp. 712–719, 2008, doi: 10.1007/978-3-540-87442-3_88.
[13] F. K. Odan and L. F. R. Reis, "Hybrid Water Demand Forecasting Model Associating Artificial Neural Network with Fourier Series," Journal of Water Resources Planning and Management, vol. 138, no. 3, pp. 245–256, 2012, doi: 10.1061/(ASCE)WR.1943-5452.0000177.
[14] Y. Wang, Y. Yuan, Y. Pan, and Z. Fan, "Modeling daily and monthly water quality indicators in a canal using a hybridwavelet-based support vector regression structure," Water (Switzerland), vol. 12, no. 5, 2020, doi: 10.3390/w12051476.