

RESEARCH NOTE

Open Access



Predicting the pulse of urban water demand: a machine learning approach to deciphering meteorological influences

Ziba Zarrin¹, Omid Hamidi^{2*}, Payam Amini³ and Zohreh Maryanaji⁴

Abstract

Objective This study delves into the impact of urban meteorological elements—specifically, air temperature, relative humidity, and atmospheric pressure—on water consumption in Kamyaran city. Data on urban water consumption, temperature (in Celsius), air pressure (in hectopascals), and relative humidity (in percent) were used for the statistical period 2017–2023. Various models, including the correlation coefficient, generalized additive models (GAM), generalized linear models (GLM), and support vector machines (SVM), were employed to scrutinize the data.

Results Water consumption increases due to the influence of relative humidity and air pressure when the temperature variable is controlled. Under specific air temperature conditions, elevated air pressure coupled with high relative humidity intensifies the response of water consumption to variations in these elements. Water consumption exhibits heightened sensitivity to high relative humidity and air pressure compared to low levels of these factors. During winter, when a western low-pressure air mass arrives and disrupts normal conditions, causing a decrease in pressure and temperature, urban water consumption also diminishes. The output from the models employed in this study holds significance for enhancing the prediction and management of water resource consumption.

Keywords Water consumption of households, Spline smoother, Simplex optimization algorithm, Nonlinear response

Introduction

Addressing the escalating challenge of urban water scarcity and exploring sustainable management methods stands out as a crucial research imperative on a global scale today [1]. Climate change and heightened global resource utilization, particularly in arid regions like Iran,

are contributing to escalating water tensions among various regions and cities. Simultaneously, the surge in urban population over recent decades has led to a doubling of the demand for water resources [2]. Access to water resources emerges as a fundamental challenge in urban management. According to studies, accurately predicting the water demand for a city relies significantly on meteorological variables, in addition to factors such as population, social dynamics, economic conditions, and technological aspects [3].

Water consumption is influenced by human factors like population and technology, but monthly and seasonal patterns depend on meteorological conditions, which drive medium-term fluctuations in consumption [4, 5]. Water scarcity in North America results from climate

*Correspondence:

Omid Hamidi
omid_hamidi@hut.ac.ir

¹Department of Photogrammetry and Remote Sensing, K. N. Toosi University of Technology, Tehran, Iran

²Department of Science, Hamedan University of Technology, Hamedan, Iran

³School of medicine, Keele University, Keele ST5 5BG, Staffordshire, UK

⁴Department of Geography, Sayyed Jamaledin Asadabadi University, Asadabad 6541835583, Iran



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

change and population growth, worsened by overexploitation of water resources. This is compounded by abundant fresh water alongside poor management practices [6].

The significance of meteorological variables in predicting urban water consumption has been underscored as crucial by various studies [7–10]. Studies have used machine learning methods like artificial neural networks (ANNs) to forecast water demand [9], specifically in relation to climatic factors. They primarily explored potential input variables, employing various statistical techniques to discern the optimal model input and to achieve precise predictions of future urban water demand, considering meteorological factors [11–13]. Seasonal climate variations influence water consumption differently throughout the year, with factors like humidity and evaporation affecting end-of-summer usage. Identifying these influences is crucial due to the complexity of urban water consumption pattern [14–16].

Effective and sustainable urban water management requires understanding meteorological variables' identification and impact. Each region has its own water consumption patterns and should be investigated by advance modelling methods to provide an insight into the effects of environmental factors. Therefore, we investigated monthly and seasonal water consumption fluctuations, using Generalized Linear Models (GLM), Generalized Additive Models (GAM), and support vector machines (SVM) in Kamyaran, Iran.

Materials and methods

Data

This study utilized data encompassing urban water consumption, temperature (measured in Celsius), air pressure (measured in hectopascals), and relative humidity (measured in percentage) for the statistical period spanning from 2017 to 2023.

Statistical analysis

The study employed zero-order Pearson correlation to explore linear associations between meteorological variables and water consumption, and first-order partial correlation to assess distinct impacts. A GAM was used to examine non-linear responses. Additionally, GLM and SVM models were constructed to capture both linear and non-linear relationships. The simplex optimizer algorithm was applied to understand the additive and subtractive influences of independent variables.

GLM extends linear models, measuring relationships between variables via regression parameters and confidence intervals, based on provided formulas. The GLM model can be written as $g(\mu) = a + \sum \beta_j x_j$ where “g” is a link function from the exponential family, μ is the

mean response, β is the vector of regression coefficients and x is the matrix of predictors.

GAMs are a nonparametric extension of GLMs [17]. GAMs allow the data to determine the shape of the response curve by $g(\mu) = a + \sum_{j=1}^p f_j(X_j)$. In these models, it is assumed that the dependent variable has a distribution from the exponential family with mean $\mu = E(Y|X_1, \dots, X_p)$, which is linked to the independent variables (X_i) through the link function (g). In fact, GAMs extend the parametric form of the independent variables in the linear model to a nonparametric form.

Here, f_j for $j = 1, 2, \dots, p$ are assumed to be unknown and smooth functions, and X_j are independent variables. Specifically, f_j is estimated from the data using advanced scatterplot smoothing techniques. GAM allows data to shape response curves unlike parametric models, replacing linear functions with smooth functions. These additive functions enable separate investigation of predictor variables' effects, identifying nonlinear relationships [18, 19].

The SVMs are data-driven algorithms, used for classification and regression problems. SVMs create a hyperplane to maximize the margin between classes, utilizing support vectors for optimization [20]. SVR is used to describe regression [21] in the form of a least squares model. In this regard, there are two types of SVMs called SVM- ϵ and SVM-Nu, and in this study, the SVM-Nu approach is used. In machine learning models, the aim is to estimate the relationship between dependent variable and independent variables, minimizing error while maintaining smoothness.

The Simplex algorithm is a non-gradient search method for minimizing functions with continuous variables. It iteratively tests solutions until an optimal point is found, making no assumptions about the function's nature [22, 23]. Therefore, in this approach, by applying the Simplex algorithm on the models produced by GLM and SVM-Nu, the simulated water consumption of each of these models was optimized.

Results and discussion

The long-term average of meteorological variables during summer for air temperature ($^{\circ}\text{C}$), relative humidity (%), and air pressure (hPa) was 8, 60, and 892 respectively. The same for summer was 29, 20, and 880 respectively. Similarly for yearly was 18, 41, and 886.

The results from the Pearson correlation analysis and the nonlinear and significant impact of temperature, pressure, and relative humidity variables at a 99% confidence level on water consumption (depicted in Table 1) reveal a statistically significant relationship between the three meteorological variables and water consumption. Specifically, air temperature exhibits a direct relationship, while air pressure and relative humidity show an inverse

Table 1 Zero- and 1-degree correlation coefficient of meteorological variables with water consumption and results of individual and combined fitting of GAM model on meteorological variables against water consumption in Khorram Abad

Correlation and Model		Air temperature (°C)	Relative humidity (%)	Air pressure (hPa)
Zero-order correlation	Water consumption	0.53	-0.50	-0.38
1-order discriminant correlation	control (air temperature)	-	0.19	0.42
	control (air pressure)	0.59	-0.32	-
	control (relative humidity)	0.42	-	0.13
Individual GAM model	Degrees of freedom	3	3	2
	Model parameters	-0.0144	-0.004	0.108
	p-value	< 0.001	< 0.001	< 0.001
Combined GAM model	Degrees of freedom	4	3.99	4
	Model parameters	0.0253	0.0049	0.0390
	p-value	< 0.001	< 0.001	< 0.001

relationship with the dependent variable. Continuing the investigation, the study explores the non-linear response of water consumption using the GAM for both individual and cumulative variables, employing the Poisson distribution and the log link function.

Figure 1 depicts the nonlinear relationship between water consumption and temperature, humidity, and pressure individually and cumulatively. Individual analysis reveals varying responses compared to the combined state, notably opposite reactions to humidity. In the cumulative state, pressure influences increased consumption, moderated by temperature. However, controlling for humidity and temperature unveils nuanced reactions, with consumption initially increasing with pressure, then decreasing beyond 883 hectopascals. A pattern of increasing-stasis-increasing is observed concerning humidity. Changes in a 5-month moving average indicate an inverse correlation between temperature and pressure, mitigating summer temperature impacts on consumption.

A 5-month moving average was applied to the time series to smooth seasonal fluctuations. Winter temperature changes have a larger impact on consumption, while summer sees mitigating effects from air pressure and humidity. Adaptive GAM and SVM models further analyze consumption reactions. Temperature increases correlate with consumption, especially in colder seasons below 11.15 °C, when humidity and air pressure are higher, indicating stronger temperature effects. (Fig. 2).

The reaction of water consumption to air temperature, pressure, and humidity in annual status, winter, and summer conditions in GLM and SVM-Nu models was checked (FIG S1). The examination of these graphs reveals a seasonal transition from winter to spring-summer, characterized by a gradual increase in air temperature and a decrease in relative humidity and air pressure. Notably, the composite spline component graphs indicate a positive impact of net pressure, relative humidity, and temperature on water consumption. Therefore, under specific conditions, the combination of these

three variables may exert a synergistic effect on water consumption.

The moving average charts depict an inverse relationship between temperature and air pressure-relative humidity trends. While rising temperatures increase water consumption, the influence diminishes in summer due to air pressure and humidity acting as mitigating factors. Consequently, the impact of temperature changes on consumption is more pronounced in cold seasons. Both SVMs (non-linear) and GAM (linear) models were employed to estimate seasonal effects on urban water consumption, revealing nonlinear reactions to meteorological variables and reducing linear correlations.

Nonlinear models generally outperform linear ones in the study. For instance, the impact of air temperature on water consumption, controlled for pressure and humidity, showed nonlinear reactions in winter and annually, but linear in summer. This highlights winter's heightened responsiveness to temperature changes due to humidity and pressure moderating summer consumption. Air pressure's effect on consumption varies, with winter conditions showing lower impact than summer. Similarly, relative humidity's impact, estimated by both linear and nonlinear models, is more pronounced in summer than winter. Overall, nonlinear models provide superior insights, indicating seasonal variations in water consumption's sensitivity to temperature, pressure, and humidity in Kamyaran's urban water management.

Discussion and conclusion

The study in Kamyaran investigates the influence of meteorological factors on urban water consumption, highlighting the interplay between temperature, pressure, and humidity. Linear and nonlinear models reveal that all three factors increase water consumption, but air pressure and humidity have a diminishing effect on the temperature-water consumption relationship. This inverse relationship is attributed to the law of gases, where temperature influences pressure and humidity. Analysis

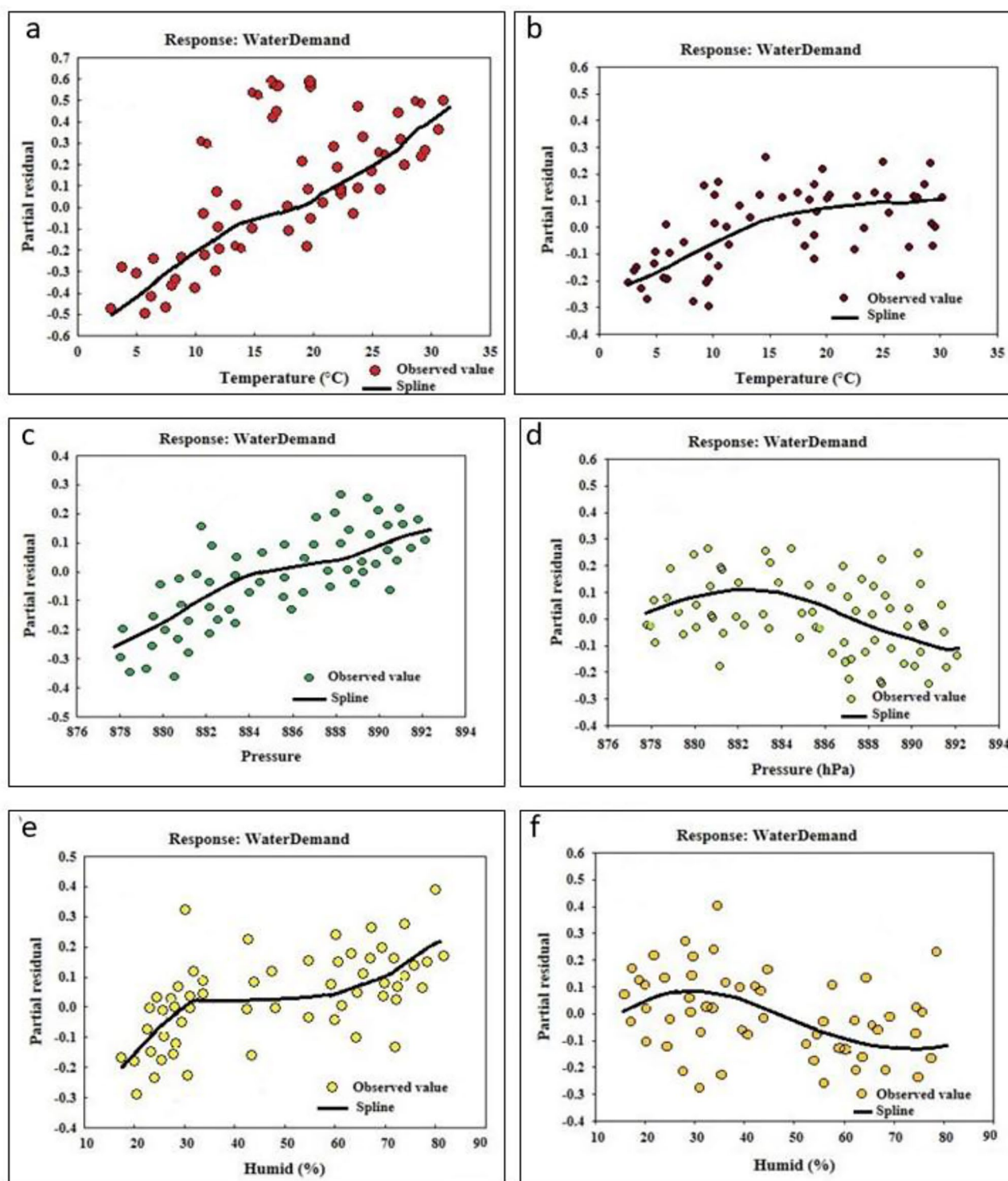


Fig. 1 (a) Cumulative reaction of water consumption to temperature, (b) Separate reaction of water consumption to temperature, (c) Cumulative reaction of water consumption to pressure, (d) Separate reaction of water consumption to pressure, (e) Cumulative reaction of water consumption to relative humidity, (f) Separate reaction of water consumption to relative humidity

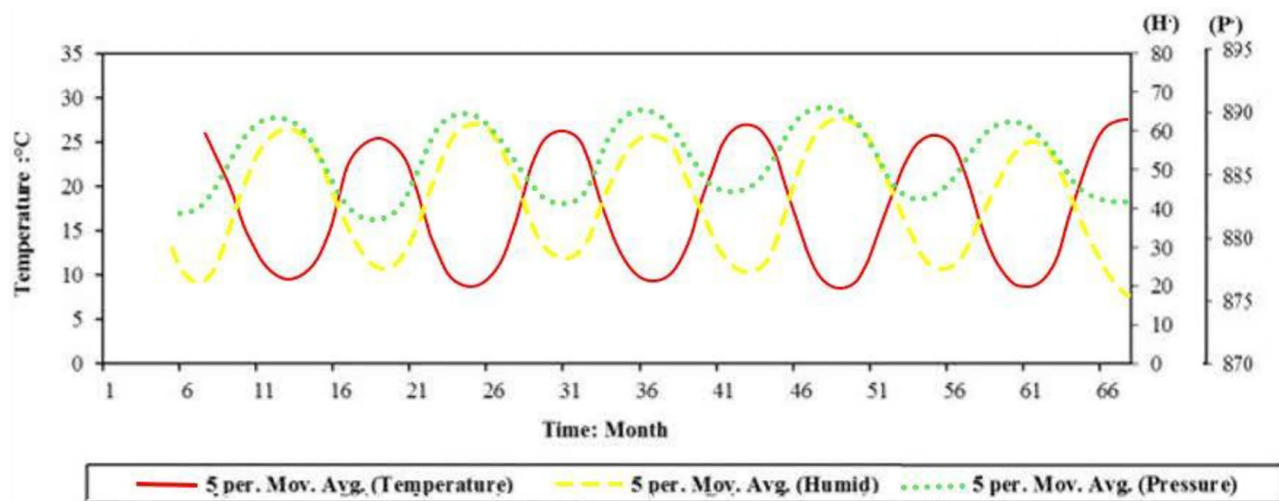


Fig. 2 Five-month moving average: temperature, pressure, relative humidity

using a 6-month moving average further elucidates these trends.

GLM and SVM models predict water consumption based on temperature, pressure, and humidity, showing stronger effects of humidity and pressure in summer and temperature in winter due to regional climate dynamics. External factors like western waves and cyclones also affect local climate, creating unique conditions where the combined effect of temperature, humidity, and pressure significantly influences water consumption, emphasizing the complexity of managing water resources in arid regions.

The absolute and relative position of the city of Kamyaran, which is influenced by the migrating western and southwestern weather systems, and with the arrival of short waves and rainy systems in the west and in the cold season, the climate of Kamyaran is subject to change. Based on the results of the partial correlation coefficient, it is expected that the entry of low-pressure air masses will have a greater effect on reducing consumption than high-pressure air masses. This is because the values of temperature and pressure in the former mass lead to a decrease in the amount of water consumption, while these two variables act against each other in the latter mass. Such results are contrary to reality and the usual and standard state of the air in Kamyaran, in which high pressure in the cold season is based on low temperatures and minimum air pressure in the hot season is based on high temperatures. Such conditions indicate the nonlinear reactions of the amount of water consumption if the effects of other variables are not kept constant. Brentan et al. [24] also showed in a study that water consumption has a positive relationship with temperature and a negative relationship with air pressure and relative humidity. Since the air pressure and relative humidity are low in the

hot season in Kamyaran, the amount of water consumption also has an upward trend.

This achievement confirms the findings by Abbasi et al. [4] who reported a threshold temperature of 15 degrees Celsius for the city of Khorramabad, which is in a similar climatic and geographical situation. Similarly, Aqeeluko and Draper [25] reported a threshold temperature of 15 degrees Celsius for the city of Calgary (Alberta, Canada), Gato et al. [26] reported a threshold temperature of 27.5 degrees Celsius for Melbourne (Australia), and Sarker et al. [27] reported a threshold temperature of 35.5 degrees Celsius for Melbourne (Australia). The research findings consistently indicate that water consumption does not exhibit a significant response to temperature increases below these identified thresholds. Maidment and Miao [28] also showed that in the states of Texas, Florida, and Pennsylvania, the reaction of water consumption to temperatures between 29 degrees Celsius and 32 degrees Celsius is about 3 to 5 times greater than temperatures below 29 degrees Celsius. Temperature thresholds were calculated individually, revealing nonlinear water consumption reactions. However, these findings lack control over other variables, showing only the individual model's effect.

Limitations

Our findings were subjected to some limitations. Shortcomings that occurred during data collection may have influenced the reliability of the results. Additionally, the small sample size was small which may limit the generalizability of our conclusions, as it may not adequately represent the broader population. Furthermore, this study does not establish causality. Without a clear causal framework, it is difficult to determine whether changes in risk factors directly influence water consumption or if other underlying factors are at play.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13104-024-06878-6>.

Supplementary Material 1

Acknowledgements

We would like to appreciate the Vice-chancellor of Education of Hamedan University of Technology for technical support.

Author contributions

Z.Z. and O.H. P.A. and Z. M. conceived the research topic, explored that idea, performed the statistical analysis and drafted the manuscript. ZZ participated in data analysis and writing. All authors read and approved the final manuscript.

Funding

This study was partially funded by Hamedan University of Technology.

Data availability

The data is available upon the request from the first author.

Declarations

Ethics approval and consent to participate

This study was approved by the Hamedan University of Technologies Ethics Committee.

Consent to publish

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 9 February 2024 / Accepted: 30 July 2024

Published online: 09 August 2024

References

- Noori A, et al. A reliable GIS-based FAHP-FTOPSIS model to prioritize urban water supply management scenarios: a case study in semi-arid climate. *Sustainable Cities Soc.* 2022;81:103846.
- Garrick D, et al. Rural water for thirsty cities: a systematic review of water reallocation from rural to urban regions. *Environ Res Lett.* 2019;14(4):043003.
- Wang X-j, et al. Adaptation to climate change impacts on water demand. *Mitig Adapt Strat Glob Change.* 2016;21:81–99.
- Hamidi O, Abbasi H, Mirhashemi H. Analysis of the Response of Urban Water Consumption to climatic variables: Case Study of Khorramabad City in Iran. *Adv Meteorol.* 2021;2021:1–14.
- Chaplot B. Prediction of rainfall time series using soft computing techniques. *Environ Monit Assess.* 2021;193(11):721.
- Asif Z et al. Climate change impacts on water resources and sustainable water management strategies in North America. *Water Resour Manage.* 2023; p. 1–16.
- Wong ST. A model on municipal water demand: a case study of northeastern Illinois. *Land Econ.* 1972;48(1):34–44.
- Sharma SK. A novel approach on water resource management with Multi-criteria optimization and Intelligent Water demand forecasting in Saudi Arabia. *Environ Res.* 2022;208:112578.
- Varahrami V. Application of genetic algorithm to neural network forecasting of short-term water demand. in *International Conference on Applied Economics–ICOAE.* 2010. Citeseer.
- Leon LP, Chaplot B, Solomon A. Water consumption forecasting using soft computing—a case study, Trinidad and Tobago. *Water Supply.* 2020;20(8):3576–84.
- Zubaidi SL, et al. Short-term urban water demand prediction considering weather factors. *Water Resour Manage.* 2018;32:4527–42.
- Polebitski AS, Palmer RN. Seasonal residential water demand forecasting for census tracts. *J Water Resour Plan Manag.* 2010;136(1):27–36.
- Chadee AA et al. Application of the analytic hierarchy process for the selection of recycling rainwater/household grey water to improve SIDS sustainability targets. *Model Earth Syst Environ.* 2023; p. 1–13.
- Kenney DS, et al. Residential water demand management: lessons from Aurora, Colorado 1. *JAWRA J Am Water Resour Association.* 2008;44(1):192–207.
- Zubaidi SL, et al. Assessing the benefits of nature-inspired algorithms for the parameterization of ANN in the prediction of Water demand. *J Water Resour Plan Manag.* 2023;149(1):04022075.
- Zamanzad-Ghavidel S, et al. Estimating of aqueduct water withdrawal via a wavelet-hybrid soft-computing approach under uniform and non-uniform climatic conditions. *Environ Dev Sustain.* 2023;25(6):5283–314.
- Hastie T, Tibshirani R. Non-parametric logistic and proportional odds regression. *J Roy Stat Soc: Ser C (Appl Stat).* 1987;36(3):260–76.
- Hastie TJ. Generalized additive models. *Statistical models in S.* Routledge; 2017. pp. 249–307.
- Vapnik V. On the uniform convergence of relative frequencies of events to their probabilities. in *Doklady Akademii Nauk USSR.* 1968.
- Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press; 2000.
- Adnan RM, et al. Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *J Hydrol.* 2020;586:124371.
- O'Neill R. Algorithm AS 47: function minimization using a simplex procedure. *J Royal Stat Soc Ser C (Applied Statistics).* 1971;20(3):338–45.
- Bishop CM. Neural networks for pattern recognition. Oxford University Press; 1995.
- Brentan BM et al. *Correlation analysis of water demand and predictive variables for short-term forecasting models* Mathematical Problems in Engineering, 2017. 2017.
- Akuoko-Asibey A, Nkemdirim L, Draper D. The impacts of climatic variables on seasonal water consumption in Calgary, Alberta. *Can Water Resour J.* 1993;18(2):107–16.
- Gato S, Jayasuriya N, Roberts P. Temperature and rainfall thresholds for base use urban water demand modelling. *J Hydrol.* 2007;337(3–4):364–76.
- Sarker R, Gato-Trinidad S, Imteaz M. *Temperature and rainfall thresholds corresponding to water consumption in Greater Melbourne, Australia.* in *20th International Congress on Modelling and Simulation.* 2013.
- Maidment DR, Miaou SP. Daily water use in nine cities. *Water Resour Res.* 1986;22(6):845–51.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.