

Water consumption prediction based on machine learning methods and public data

Witwisit Kesornsit^{1a} and Yaowarat Sirisathitkul^{*2}

¹Government Data Solution Division, Department of Data Solution, Digital Government Development Agency (Public Organization), Bangkok, Thailand

²Department of Computer Engineering and Electronics, School of Engineering and Technology, Walailak University, Nakhon Si Thammarat, Thailand

(Received September 16, 2021, Revised December 14, 2021, Accepted December 26, 2021)

Abstract. Water consumption is strongly affected by numerous factors, such as population, climatic, geographic, and socio-economic factors. Therefore, the implementation of a reliable predictive model of water consumption pattern is a challenging task. This study investigates the performance of predictive models based on multi-layer perceptron (MLP), multiple linear regression (MLR), and support vector regression (SVR). To understand the significant factors affecting water consumption, the stepwise regression (SW) procedure is used in MLR to obtain suitable variables. Then, this study also implements three predictive models based on these significant variables (e.g., SWMLR, SWMLP, and SWSVR). Annual data of water consumption in Thailand during 2006 – 2015 were compiled and categorized by provinces and distributors. By comparing the predictive performance of models with all variables, the results demonstrate that the MLP models outperformed the MLR and SVR models. As compared to the models with selected variables, the predictive capability of SWMLP was superior to SWMLR and SWSVR. Therefore, the SWMLP still provided satisfactory results with the minimum number of explanatory variables which in turn reduced the computation time and other resources required while performing the predictive task. It can be concluded that the MLP exhibited the best result and can be utilized as a reliable water demand predictive model for both of all variables and selected variables cases. These findings support important implications and serve as a feasible water consumption predictive model and can be used for water resources management to produce sufficient tap water to meet the demand in each province of Thailand.

Keywords: artificial neural network; machine learning; multi-layer perceptron; multiple linear regression; predictive model; stepwise regression; support vector regression; water consumption

1. Introduction

Over the past decades, fast-growing demand for water has been a global concern, especially in terms of household, agriculture, and industrial undertaking. Moreover, water provides a link between different elements in an ecosystem (Stampoulis *et al.* 2021). Many parts of the world are facing the risks of severe water scarcity situations (Yang *et al.* 2021). Water shortage has become a

*Corresponding author, Assistant Professor, E-mail: kinsywu@gmail.com; syaowara@mail.wu.ac.th

^aM.Sc. Research Scholar., E-mail: witwisit.kes@gmail.com

- Wei, H.W., Hassan, M., Che, Y., Peng, Q.K., Wang, Q., Su, Y.L. and Xie, B. (2021), “Spatio-temporal characteristics and source apportionment of water pollutants in upper reaches of Maotiao River, Southwest of China, from 2003 to 2015”, *J. Environ. Inform.*, **37**(2), 93-106. <http://doi.org/10.3808/jei.201900415>.
- Wongso, E., Nateghi, R., Zaitchik, B., Quiring, S. and Kumar, R. (2020), “A data-driven framework to characterize state-level water use in the United States”, *Water Resour. Res.*, **56**(9), e2019WR024894. <https://doi.org/10.1029/2019WR024894>.
- Wu, L. and Zhou, H. (2010), “Urban water demand forecasting based on HP filter and fuzzy neural network”, *J. Hydroinformatics*, **12**(2), 172-184. <https://doi.org/10.2166/hydro.2009.082>.
- Xiao, Y., Fang, L. and Hipel, K.W. (2021), “Conservation-targeted hydrologic-economic models for water demand management”, *J. Environ. Inform.*, **37**(1), 49-61. <http://doi.org/10.3808/jei.201900418>.
- Yang, Y., Huang, T.T., Shi, Y.Z., Wendroth, O. and Liu, B.Y. (2021), “Comparing the performance of an autoregressive state-space approach to the linear regression and artificial neural network for streamflow estimation”, *J. Environ. Inform.*, **37**(1), 36-48. <https://doi.org/10.3808/jei.202000440>.
- Zhong, H., Wang, J., Jia, H., Mu, Y. and Lv, S. (2019), “Vector field-based support vector regression for building energy consumption prediction”, *Appl. Energy*, **242**, 403-414. <https://doi.org/10.1016/j.apenergy.2019.03.078>.
- Zubaidi, S.L., Dooley, J., Al-Khaddar, R., Abdellatif, M., Al-Bugharbee, H. and Ortega-Martorell, S. (2018), “A novel approach for predicting monthly water demand by combining singular spectrum analysis with neural networks”, *J. Hydrol.*, **561**, 136-145. <https://doi.org/10.1016/j.jhydrol.2018.03.047>.
- Zubaidi, S.L., Ortega-Martorell, S., Al-Bugharbee, H., Olier, I., Hashim, K.S., Gharghan, S.K., Kot, P. and Al-Khaddar, R. (2020a), “Urban water demand prediction for a city that suffers from climate change and population growth: Gauteng province case study”, *Water*, **12**(7), 1885. <https://doi.org/10.3390/w12071885>.
- Zubaidi, S.L., Ortega-Martorell, S., Kot, P., Al-Khaddar, R., Abdellatif, M., Gharghan, S.K., Ahmed, M.S. and Hashim, K.S. (2020b), “A method for predicting long-term municipal water demands under climate change”, *Water Resour. Manag.*, **34**(3), 1265-1279. <https://doi.org/10.1007/s11269-020-02500-z>.