APPLICATION OF BOOSTED CLASSIFICATION AND REGRESSION MODELS FOR GROUNDWATER POTENTIAL MAPPING

Sunmin Lee (1), (2), Moung-Jin Lee (2)*, Jeong-Cheol Kim (1), (3)

[1] University of Seoul, 163 Seoulsiripdaero, Dongdaemun-gu, Seoul, 02504, Korea
[2] Center for Environmental Assessment Monitoring, Korea Environment Institute (KEI), 370 Sicheong-daero, Sejong-si, 30147, Korea
[3] Riparian Ecosystem Research Team, National Institute of Ecology (NIE), 1210, Geumgang-ro, Chungcheongnam-do, 33657, Korea
Email: smilee@kei.re.kr; leemj@kei.re.kr; jckim@nie.re.kr

KEY WORDS: groundwater potential; boosted tree; specific capacity

ABSTRACT: An appropriate management of groundwater in rural area is significant because it is responsible for both agricultural and drinking water functions. Therefore, in this study, Boosted Classification and Regression Models applied for groundwater potential mapping in Yangpyeong, rural area near Seoul city. Total 19 groundwater related factors were applied for mapping including topographical factors, hydrological factors, forest, soil, land use and geological factors. 53 well locations of specific capacity data were collected and the data were randomly used for the training with 70% and for the validation with 30% of the well location data, respectively. The models of Boosted Classification and Regression and Tree were applied to produce the groundwater potential maps. Validation for these maps was executed by using receiver operational characteristics (ROC) method. The area under the ROC showed 82.54 % and 79.09 % of accuracy for the classification and regression for boosted tree models, respectively. As a result, boosted tree models which are used in this study could effectively predict the groundwater potential in data-scarce areas, especially for rural area where groundwater is considered relatively important. The results from in this study could be used for the rational and continuous management of groundwater.

1. INTRODUCTION

Globally, groundwater consumption is increasing due to population growth, urbanization and industrialization (Zahedi et al., 2017). As the demand for groundwater increases, various researches are conducted actively to develop quantitative methodologies for evaluating groundwater potential in aquifers (Golkarian et al., 2018). Groundwater potential is affected by various geoenvironmental factors such as topography, topography, geological structure, land use and climate (Dušek and Velísková, 2017; Shenga et al., 2018). Correlation between these factors therefore plays an important role in the identification and quantification of groundwater resources in specific regions (Golkarian and Rahmati, 2018). Therefore, to develop an accurate groundwater potential model, it is important to determine the relationship between the set of influence factors and groundwater location.
For the acquisition of groundwater related data, remote sensing for exploration of groundwater potential zones have been carried out by many researchers around the world, and the results vary due to the different factors involved in determining groundwater potential zones. Groundwater exploration and merged other factors such as drainage density, topography, geology, slope, land use, rainfall intensity, and soil organization were represented (Sander et al., 1996). The results obtained were found to be satisfactory according to field surveys and vary from region to region due to various geographic environmental conditions. These techniques have surpassed traditional geophysical, geological and hydrogeological approaches (Rahmati et al., 2016). Therefore, in this study, boosted tree models of classification and regression applied for groundwater potential mapping in Yangpyeong, rural area.

2. DATA AND METHODOLOGY

2.1 Study area and data

Yangpyeong-gun is located in the northeastern part of Gyeonggi-do in the midwest of the Korean peninsula. It is adjacent to Hongcheon-gun, Gangwon-do in the northeast, Hoengseong-gun in the east, Wonju-si in Gangwon-do in the southeast, Yeoju-si, Gyeonggi-do in the south, Gwangju-si, Gyeonggi-do in the southwest, Namyangju-si in Gyeonggi-do, and Gapyeong-gun in Gyeonggi-do in the north. The hydraulic and latitudinal positions occupied by Yangpyeong-gun are 127º18′46 ″-127º51′02 ″ east longitude and 37º21′33″-37º40′07″ north latitude.

Various groundwater related factors were used for groundwater potential modeling in this study (Table 1). Groundwater well data for pumped groundwater specific capacity from 53 wells in the study area was collected and used for training and validation sets. Topographic, geological, hydrological and surface area factors are commonly applied to predict groundwater potential. Depending on the local characteristics, it is necessary to take into account the conditioning factors. For this reason, this study estimated groundwater potential using 19 factors. Total 19 groundwater related factors were applied for mapping including topographical factors, hydrological factors, forest, soil, land use and geological factors. The related factor is prepared with 30m spatial resolution in GIS environment.

2.2 Methodology

The boost tree model is an ensemble machine learning model of decision trees. Like other ensemble models based on decision trees, boost tree models fit multiple decision trees to improve the accuracy of results based on boosting methods. The boost tree model iteratively selects a random subset for each new tree from the entire data set of identical data (De'Ath, 2007).

The difference between other models and boost trees is how you select a random subset. The boost tree model uses a sequential approach for random subset sampling. It uses the boosting method of weighting the incorrectly modeled data from the old tree to select for the new tree. Once the subsequent tree is selected, the model will try to select the next tree, taking into account the prediction errors of the previous tree. Therefore, after mounting the first tree, the model can continuously improve its accuracy, taking into account the suitability of the previous tree (Elith et al., 2008).

3. RESULT AND CONCLUSION

Figure 1 shows the groundwater potential maps that resulted from this study. Subsequent verifications were based on 30% of the well points of the groundwater potential model performed.

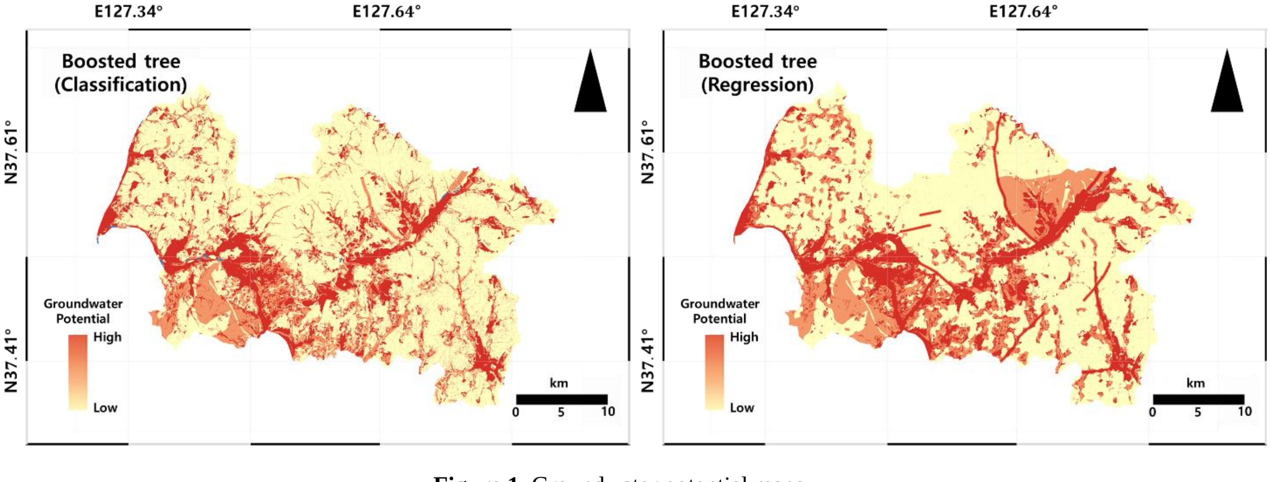

Figure 1. Groundwater potential maps.

Among the boost tree results, the classification model is 82.54 % and the regression model is 79.09 %, as shown in Fig. 2. It should be noted that both models used in this study have achieved acceptable accuracy for groundwater potential mapping.

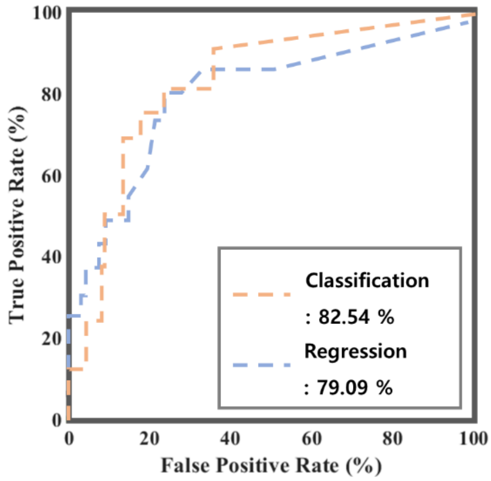

Figure 2. Validation Results**.**

This result can be applied to groundwater exploration and other areas that require groundwater utilization of groundwater recharge. Managers and policy makers can effectively analyze groundwater potential modeling results to maximize their benefits.

REFERENCES

De'Ath, G. (2007). Boosted trees for ecological modeling and prediction. Ecology, 88(1), 243-251.

Dušek, P., & Velísková, Y. (2017). Comparison of the MODFLOW modules for the simulation of the river type boundary condition. Pollack Periodica, 12(3), 3-13.

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. Journal of Animal Ecology, 77(4), 802-813.

Golkarian, A., & Rahmati, O. (2018). Use of a maximum entropy model to identify the key factors that influence groundwater availability on the Gonabad Plain, Iran. Environmental earth sciences, 77(10), 369.

Golkarian, A., Naghibi, S. A., Kalantar, B., & Pradhan, B. (2018). Groundwater potential mapping using C5. 0, random forest, and multivariate adaptive regression spline models in GIS. Environmental monitoring and assessment, 190(3), 149.

Rahmati, O., Pourghasemi, H. R., & Melesse, A. M. (2016). Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran Region, Iran. Catena, 137, 360-372.

Sander, P., Chesley, M. M., & Minor, T. B. (1996). Groundwater assessment using remote sensing and GIS in a rural groundwater project in Ghana: lessons learned. Hydrogeology Journal, 4(3), 40-49.

Shenga, Z. D., Baroková, D., & Šoltész, A. (2018). Modeling of groundwater extraction from wells to control excessive water levels. Pollack Periodica, 13(1), 125-136.

Zahedi, S., Azarnivand, A., & Chitsaz, N. (2017). Groundwater quality classification derivation using multi-criteria-decision-making techniques. Ecological Indicators, 78, 243-252.