Government of Nepal
Ministry of Energy, Water Resources and Irrigation
Water Resources Research and Development Centre
Pulchowk, Lalitpur

WRRDC RESEARCH LETTER

Issue: 17
February 2022

Delineation of groundwater potential zones in a slope unit-scale by means of naïve Bayes classifier

Background

This research presents a machine learning method to map groundwater potential in hardrock domains. First, a spatially-distributed set of explanatory variables for groundwater occurrence is complied into a geographic information system. Groundwater potential may be inferred from existing maps (lithological cartographies or soil type), digital elevation models–and derived products, aerial photograph interpretation, satellite imagery and geophysical information. Traditional groundwater potential studies were exclusively based on expert judgment techniques, including analytical hierarchy and weight of evidence processes. These methods require a grouping of the variables in intervals. The advent of machine learning (ML), i.e., artificial intelligence systems apt to learn without explicit instructions by using algorithms and statistical models, opens up a new methodological dimension to groundwater studies.

Khandbari Municipality lies in the Sankhuwasabha district which is the fastest growing city and is most developed city among hilly region. It is also the main departure point for the 5th highest mountain, Makalu. Sankhuwasabha district is identified as the most vulnerable in case of water having the low coverage and degree of inequality. The people of Khandbari depends upon the spring water for drinking, washing, bathing processes.

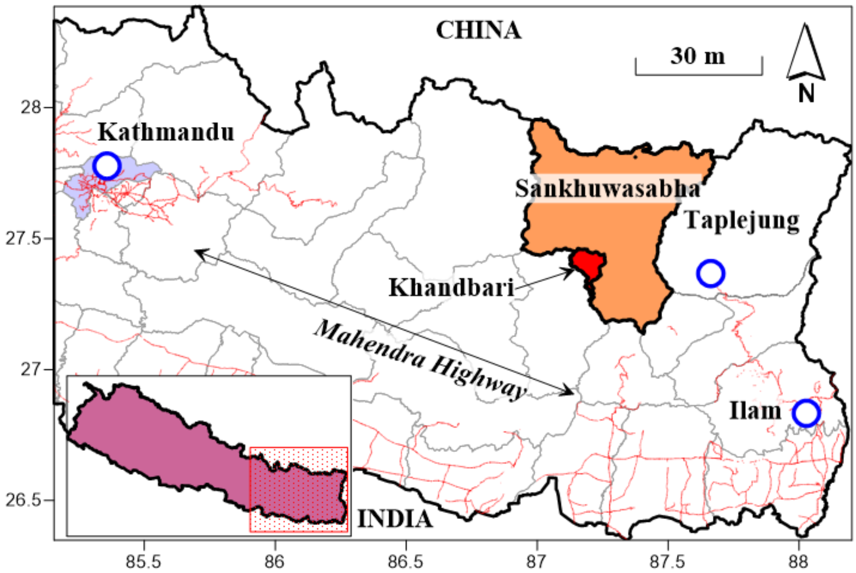

Figure 1: Location of study area

Figure 2: Observed springs

Method

Naïve Bayes (NB) is based on Bayes' theorem and, in this case, makes the naïve assumption of strong independence among groundwater factors. We defined x = (x_1, x_2, …, x_n) as a vector of the CFs, y = (y_1, y_2) is a vector of the classifier dependent variables (spring, non-spring). The Bayes' theorem is shown in the following relationship:

$$y_{NBc} = \underset{y_j=[spring,non-spring]}{\arg\max}\, p(y_j)\prod_{i=1}^{n} p(x_i \mid y_j) \quad \text{.......(1)}$$

$$p(x_i \mid y_j) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{\frac{-(x_i-\mu_i)^2}{2\sigma_i^2}} \quad \text{.......(2)}$$

where y_{NBc} is the groundwater probability, $p(y_j)$ is the prior probability of y_j, and $p(x_i|y_j)$ is the conditional probability that can be calculated from the Gaussian distribution. μ_i is the mean and σ_i is the standard deviation of x_i.

We choose slope units as the basic unit type for the groundwater zonation. For the analysis methods, we focused mainly on the application of existing mature models to produce more groundwater potential zonation, rather than exploring and trying the latest models.

Compared with traditional grid cells, slope units are able to better reflect the actual environmental conditions that lead to accumulate groundwater and have definite geological significance. In this work, the hydrological analysis module of ArcGIS was used to produce the slope units for the groundwater potential zonation from SRTM 30 m digital elevation model (DEM) data. A total of 2842 slope units were obtained, of which the minimum area was 0.05 m^2 and the maximum area was 0.47 km^2.

Results

In the case at hand, the target variable is groundwater potential, which is defined in binary terms as the likelihood that a spring at a given location would be positive. The variables include lithology, geological lineaments, landforms, topography, land cover, drainage, slope-related variables, rainfall, and vegetation indices, among others. Groundwater potential zonation assumes that the presence or absence of groundwater can be partially inferred from surface features. This process includes collinearity, cross-validation, feature elimination and parameter fitting routines. Naïve Bayes classifiers outperformed other algorithms. In the initial phase, the database was divided into 70% spring points for training and 30% for validation (testing) of model performance.

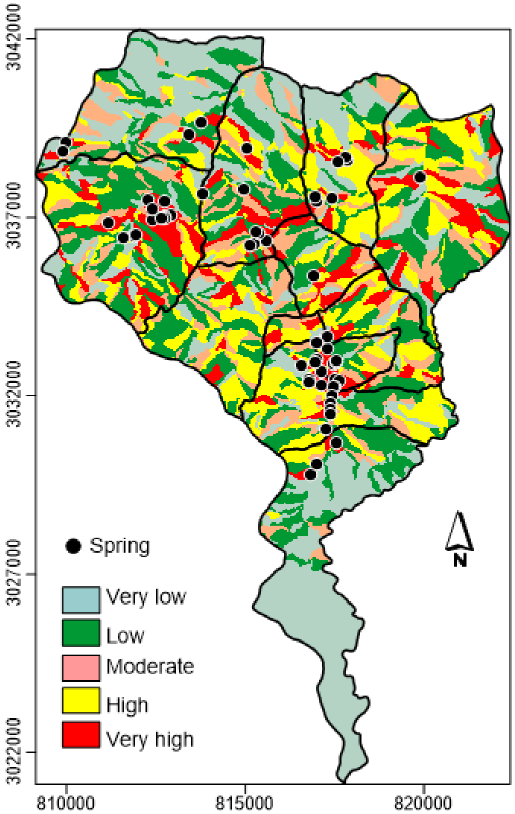

Figure 3: Groundwater potential zone

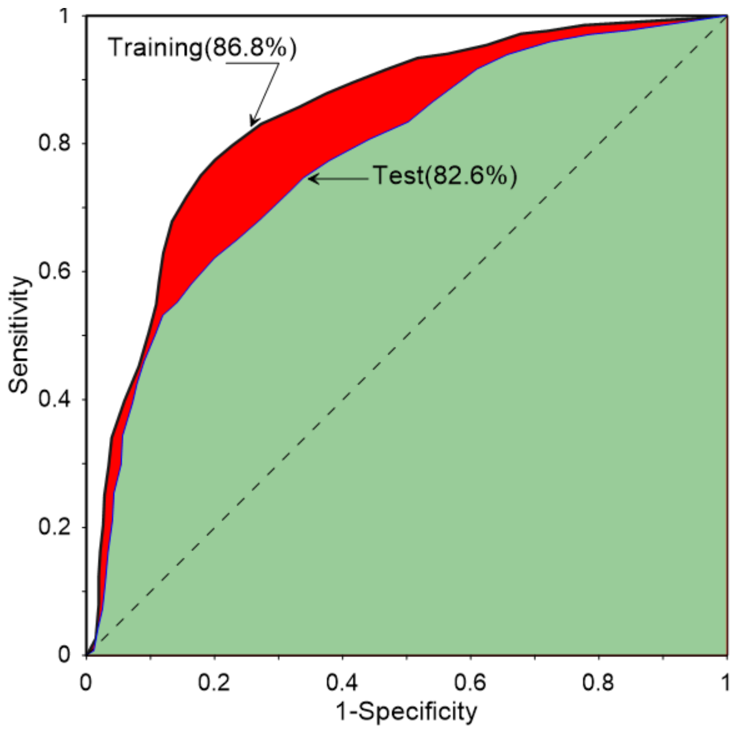

Figure 4: Accuracy of the model

Remarks

The result demonstrated the capability of Naïve Bayes classifiers for the prediction of groundwater potential zones in the mountainous terrain. We can apply a similar methodology in other mountainous terrain with similar geo-environmental conditions.

FOR FURTHER INFORMATION:
Government of Nepal
Ministry of Energy, Water Resources and Irrigation
Water Resources Research and Development Centre
Dr. Ananta Man Singh Pradhan, Sr.Div. Engineering Geologist, ananta.pradhan@nepal.gov.np
Mahesh Khanal, Information Officer, mahesh.khanal@nepal.gov.np
Website: www.wrrdc.gov.np

Citation: Pradhan, A.M.S., Shrestha, A., 2022. Delineation of groundwater potential zones in a slope unit-scale by means of naïve Bayes classifier. WRRDC Research Letter, Issue no: 17, February 2022, pp. 1-2.