# Spatial Prediction of Groundwater Spring Potential Mapping Based on Adaptive Neuro-Fuzzy Inference System and Metaheuristic Optimization

3

Khabat Khosravi<sup>1</sup>, Mahdi Panahi<sup>\*2</sup>, Dieu Tien Bui<sup>\*3</sup>

4 1-Department of watershed management engineering, Faculty of Natural Resources, Sari Agricultural Science and

5 Natural Resources University, Sari, Iran. (E-mail: khabat.khosravi@gmail.com)

6 2-Young Researchers and Elites Club, North Tehran Branch, Islamic Azad University, Tehran, Iran. (E-mail:

7 panahi2012@yahoo.com)

8 3- GIS group, Department of Business and IT, University of South-Eastern Norway, Gullbringvegen 36, 3800 Bø i

9 Telemark, Norway. (E-mail: Dieu.T.Bui@usn.no)

10 11

# 12 Abstract

Groundwater is one of the most valuable natural resources in the world; however, the 13 groundwater is not an unlimited resource, therefore understanding groundwater potential is 14 15 crucial to ensure its sustainable use. The aim of the current study is to propose and verify new artificial intelligence methods for spatial prediction of groundwater spring potential mapping at 16 the Koohdasht-Nourabad plain, Lorestan province, Iran. These methods are new hybrids of 17 Adaptive Neuro-Fuzzy Inference System (ANFIS) and five meta-heuristic algorithms, namely 18 Invasive Weed Optimization (IWO), Differential Evolution (DE), Firefly Algorithm (FA), 19 Particle Swarm Optimization (PSO), and Bees Algorithm (BA). A total of 2463 spring locations 20 were identified and collected, and then, divided in two subsets randomly: 70% (1725 locations) 21 were used for training models and the remaining (30%, 738 spring locations) were utilized for 22 23 evaluating the models. Thirteen groundwater conditioning factors were prepared for modeling, namely slope degree, slope aspect, altitude, plan curvature, stream power index (SPI), 24 topographic wetness index (TWI), terrain roughness index (TRI), distance from fault, distance 25 26 from river, land-use/land-cover, rainfall, soil order, and lithology. In the next step, the Stepwise Assessment Ratio Analysis (SWARA) method was applied to quantify the degree of relevance of 27 these groundwater conditioning factors. The global performance of these derived models was 28 29 assessed using the Area Under the Curve (AUC). In addition, the Friedman and Wilcoxon signed rank tests were carried out to check and confirm the best model to use in this study. The result 30 showed that all models have high prediction performance; however, the ANFIS-DE model has 31 32 the highest prediction capability (AUC = 0.875), followed by the ANFIS-IWO model, the 33 ANFIS-FA model (0.873), the ANFIS-PSO model (0.865), and the ANFIS-BA model (0.839). The results of this research can be useful for decision makers responsible for sustainable 34 35 management of groundwater resources.

36

Keywords: Groundwater spring, ANFIS-DE, ANFIS-IWO, ANFIS-FA, ANFIS-PSO, ANFIS-BA, Iran.

- 39
- 40

#### 42 **1. Introduction**

Groundwater is defined as the water in a saturated zone which fills rock and pore spaces 43 (Berhanu et al., 2014; Fitts, 2002), whereas groundwater potential is the probability of 44 groundwater occurrence in an area (Jha et al., 2010). The occurrence and movement of 45 groundwater in an aquifer are affected by various geo-environmental factors including lithology, 46 topography, geology, fault and fracture and its connectivity, drainage pattern, and land-use/land-47 cover (Mukherjee, 1996). Geological strata acts like a conduit and reservoir for groundwater 48 while storage and transmissivity influence the suitability of exploitation of groundwater in a 49 given geological formation. Downhill and depression slopes impart runoff and improve recharge 50 and infiltration, respectively (Waikar and Nilawar, 2014). Globally, groundwater is a major 51 source of drinking water for around two billion people(Richey et al., 2015), whereas in 52 agriculture, about 278.8 million ha of agricultural lands are irrigated by the groundwater (Siebert 53 et al., 2013). Due to population and economic growth, the demand of groundwater is anticipated 54 increasing in the future (Ercin and Hoekstra, 2014). For the case of Iran, approximately two-third 55 of the land is covered by deserts and groundwater is still the main water source for drinking and 56 57 other uses (Nosrati and Van Den Eeckhaut, 2012). According to Rahmati et al. (2016), groundwater in Iran supplies around 65% of the water use-up and the remaining 35% is supplied 58 by surface water. However, groundwater is not an unlimited resource, therefore understanding 59 groundwater potential is crucial to ensure its sustainable use. One of the most efficient methods 60 for the protection and management of groundwater to identify groundwater potential zoning 61 (Ozdemir, 2011b). 62

There are a number of methods for groundwater potential zoning and exploitation. Traditional methods i.e. drilling, geological, geophysical, and hydrogeological methods are the most widely used (Israil et al., 2006; Jha et al., 2010; David Keith Todd and Mays, 1980; Sander et al., 1996; Singh and Prakash, 2002). However, they are time-consuming and costly methods, especially for large areas. In recent years, Geographic Information Systems (GIS) and remote sensing (RS) have become effective tools for groundwater potential mapping (Fashae et al., 2014) due to their ability in handling huge amount of spatial data.

70 In more recent years, some probabilistic models such as frequency ratio (Oh et al., 2011), multicriteria decision analysis (Kaliraj et al., 2014; Rahmati et al., 2015), weights-of-evidence (WoE) 71 (Pourtaghi and Pourghasemi, 2014), logistic regression (Ozdemir, 2011a; Pourtaghi and 72 Pourghasemi, 2014), evidential belief function (Nampak et al., 2014; Pourghasemi and 73 74 Beheshtirad, 2015), and Shannon's entropy (Naghibi et al., 2015) have been considered for groundwater potential mapping. Bivariate and multivariate statistical models have disadvantages 75 76 in measuring the relationship between groundwater occurrence and conditioning factors (Tehrany et al., 2013; Umar et al., 2014), whereas MCDA technique is a source of bias due to 77 expert opinion. Therefore, machine learning has been considered and has proven efficient due to 78 ability to handle non-linear structured data from various sources with different scales. In 79 addition, machine learning requires no statistical assumptions. Among machine learning 80 methods, artificial neural network (ANN) is a widely used method for groundwater mapping due 81 to its computational efficiency (Sun et al., 2016; Mohanty et al., 2015; Maiti and Tiwari, 2014). 82 83 However, the ANN model has a number of weaknesses such as poor prediction and error in modeling process (Bui et al., 2016); therefore, hybrid models have been proposed. Among 84 hybrid frameworks, ensemble of fuzzy logic and neural networks i.e. Adaptive Neuro-Fuzzy 85

Inference System (ANFIS) has proven it's efficient in term of high accuracy (Lohani et al., 2012; Emamgholizadeh et al., 2014; Zare and Koch, 2018; Nourani et al., 2016). It should be noted that although ANFIS model has a higher accuracy than other models, it is still difficult in finding the best internal weight values of ANFIS due to the limited nature of the adaptive algorithm used (Bui et al., 2016). Thus, these weights should be optimized by new metaheuristic optimization algorithms to enhance the prediction accuracy of ground water models.

92 The main goal of the current study is to propose and verify integration of new metaheuristic 93 optimization algorithms with ANFIS for groundwater spring potential mapping (GSPM) in 94 Koohdasht-Nourabad plain, Iran. Accordingly, five new metaheuristic algorithms are 95 investigated, Invasive Weed Optimization (IWO), Differential Evolution (DE), Firefly Algorithm 96 (FA), Particle Swarm Optimization (PSO), and Bees algorithm (BA). According to current 97 literature, it is the first time such study is conducted for groundwater potential mapping.

#### 98 **2. Description of the study area**

99 Koohdasht-Nourabad plain is located in the western part of the Lorestan province (Iran) and covers an area of around 9531.9 km<sup>2</sup>. It lies between latitudes 33°3' 28 N and 34° 22' 55 N, and 100 between longitudes 46° 50' 19 E and 48° 21' 18 E (Figure 1). The region is located in the semi-101 arid area with mean annual precipitation of about 450 mm (Iran Meteorological Organization). 102 The altitude of the study area varies between 531 m and 3175 m above sea level. The maximum 103 slope and minimum slope are 64° and 0°, respectively. Geologically, the study area is located in 104 105 the Zagros structural zone of Iran and is mostly covered by Quaternary and Cretaceous-Paleocene geologic time scale. The dominant land-use/land-cover of the study area is moderate 106 forest (20%). The residential areas cover about 3% of the plain. Rock crop/inceptisoils are the 107 dominant soil types in the study area and covering about 51% of the study area. Population of the 108 plain is 362,000 people (according to 2016 census) and agriculture is the primary occupation. In 109 this plain, groundwater is the main water source for drinking and agricultural activities. 110



Figure 1. Groundwater well locations with DEM of the study area.

**3. Methodology** 





**Figure 2**. Conceptual model of methodology applied in the current study.

#### 118 **3.1. Data preparation**

#### 119 **3.1.1. Groundwater spring inventory map**

In groundwater modeling, spatial relationships between groundwater springs and conditioning factors should be analyzed and assessed to determine the best subset of these factors. In the Koohdasht-Nourabad plain, a total of 2463 spring locations were provided by Iranian Water Resources Management Bureau. Most of these spring locations were checked during extensive field surveys using a GPS hand held unit.

### 125 **3.1.2.** Construction of the training and testing datasets

Spatial prediction of groundwater potential mapping using machine learning model is considered a binary classification with two classes, spring and non-spring. Therefore, a total of 2463 nonspring locations were randomly generated using the random point tool in ArcGIS10.2. According to Chung and Fabbri (2003), it is possible to validate the model performance using a cross validation method that splits the dataset into the two parts for both of spring and non-spring location. The first part is used for model building which is called a training dataset and the other part is utilized for validating the model performance named as a testing dataset (Pham et al., 2017a). In this study, a ratio of 70/30 was selected randomly for generating the training and testing the dataset (Pourghasemi et al., 2012; Pourghasemi et al., 2013a; Pourghasemi et al., 2013b; Xu et al., 2012). Accordingly, both spring locations and non-spring locations have been randomly divided into two groups for training (1725 locations) and validating (738 locations) purposes (Figure 1).

Both the training and the testing datasets were converted to raster format where spring pixels were assigned as "1" and non-spring pixels were assigned to "0" (Bui et al., 2015), and then, these pixels were overlaid with 13 groundwater conditioning factors to extract their attribute values.

# 142 **3.1.3.** Groundwater conditioning factor analysis

# 143 **3.1.3.1. Selection of the Groundwater conditioning factor**

After the initial selection of the conditioning factors, these factors should be assessed for multi-144 collinearity problems. Multi-collinearity takes place when two or more independent conditioning 145 146 factors are highly correlated or in other words inter-dependent (Li et al., 2010). Several methods have been proposed to diagnose multi-collinearity, and among them, Variance Inflation Factor 147 (VIF) and Tolerance (TOL) are widely used in environmental modeling (O'brien, 2007; Bui et 148 al., 2016); therefore, they were selected for this research. Factors with VIFs greater than 5 and 149 TOL are less than 0.1 indicate multi-collinearity problems existed (O'brien, 2007; Bui et al., 150 2011). Another method namely Information Gain Ratio (IGR) technique was applied to 151 identifying the relative importance of the conditioning factor and also to obtain factors with null 152 effect. These factors must be removed to increase the accuracy of the model (Khosravi et al, 153 2018). 154

155 In the current study, 13 conditioning factors have been selected, namely slope degree, slope aspect, altitude, plan curvature, stream power index (SPI), topographic wetness index (TWI), 156 Terrain roughness index (TRI), distance from fault, distance from river, land-use/land-cover, 157 rainfall, soil order, and lithology units. These factors have been determined based on literature 158 review, characteristics of the study area, and data availability (Nampak et al., 2014; Mukherjee, 159 1996; Oh et al., 2011; Ozdemir, 2011b). The process of converting continuous variables into 160 161 categorical classes were carried out based on our frequency analysis of springs location (Khosravi et al, 2018) in order to define the class intervals. 162

Digital Elevation Model (DEM) for the study area was downloaded from ASTER global DEM 163 (https://asterweb.jpl.nasa.gov/gdem.asp) with 30x30 m grid size. Based on the DEM, slope 164 degree, slope aspect, altitude, plan curvature, SPI, TWI and TRI were derived. Slope degree has 165 been divided in five categories using the quantile classification scheme (Tehrany et al., 2013, 166 2014), namely 0-5.5, 5.5-12.11, 12.11-19.4, 19.4-28.7, 28.7-64.3 degrees (Figure 3a). Slope 167 aspect is selected because it controls solar radiation budgets that influence the groundwater 168 potential. Slope aspect has been provided in 5 different classes, flat, north, west, south and east 169 (Figure 3b). Altitude was divided into five classes using the quantile classification scheme, 170 namely 531-1070, 1070-1385, 1385-1703, 1703-2068 and 2068-3175 m (Figure 3c). Plan 171 curvature was divided into three classes, namely concave (<-0.05), flat (-0.05-0.05), and 172 convex (>0.05) (Figure 3d) (Pham et al. 2017). 173

SPI is related to erosive power of surface runoff, whereas TWI relates to amount of the flow that
accumulates at any point in the catchment. In this research, SPI, TWI and TRI were constructed
using the System for Automated Geoscientific Analyses SAGA-GIS 2.2 software, and finally,
were divided into five classes. These classes are 0-48664, 48664-227099, 227099-583969,
583969-1330153, and 1330153-4136452 for SPI (Figure 3e). For TWI, these classes are 2.1-4.6,
4.6-5.6, 5.6-6.6, 6.6-7.9, 7.9-11.9 (Figure 3f), and for TRI, these classes are 0-8.7, 8.7-18.2, 18.229.9, 29.9-46.6, 46.6-185 (Figure 3g).

Distance from fault and river for the study area were generated with five classes using the 181 multiple ring-buffer tool in ArcGIS10.2, 0-200, 200-500, 500-1000, 1000-2000 and >2000 m 182 (Figures 3h and 3i). Lithology plays a key role in determining the groundwater potential due to 183 different infiltration rate of formation (Adiat et al., 2012; Nampak et al., 2014). Land-use/land-184 cover of the study area was obtained using Landsat 7 Enhanced Thematic Mapper plus (ETM+) 185 downloaded from the US Geological Survey (available 186 images that are at https://earthexplorer.usgs.gov). Accordingly, 25 land-use/land-cover types were recognized: 187 agriculture (A), garden (G), dense-forest (DF), good rangeland (GR), poor forest (PF), waterway 188 (W), mixture of garden and agriculture (MGA), mixture of agriculture with dry farming 189 (MADF), mixture of agriculture with poor-garden (MAPG), dry farming (DF), follow (F), dense 190 rangeland (DR), very poor forest (VPF), mixture of waterway and vegetation (MWV), mixture of 191 moderate forest and agriculture (MMFA), mixture of moderate rangeland and agriculture 192 (MMRA), mixture of poor rangeland and follow (MPRF), mixture of low forest and follow 193 (MLFF), wood-land (WL), moderate forest (MF), moderate rangeland (MR), poor rangeland 194 195 (PR), bare soil and rock (BSR), urban and residential (UR), mixture of very poor forest (MVPF), and rangeland (R) have been identified (Figure 3i). 196

Rainfall is the major source of recharge to the groundwater. In this research, mean annual rainfall data of 15 years (2000–2015) at 4 rain-gauge stations of the study area was used. The rainfall map (Figure 3k) with five categories (300-400, 400-500, 500-600, 600-700, and 700-800 mm) was generated using Inverse distance weighted method due to lower RMSE (Khosravi et al, 2016a,b). Soil map at scale of 1:50,000 for the study area was provided by the Iranian Water Resources Department (IWRD). The soil types are soil rock outcrop/entisols, rock outcrop/inceptisols, inceptisols, inceptisols, and badlands (Figure 3l).



Figure 3. Groundwater conditioning factors for the study area used in this research: (a) slope degree; (b)
slope aspect; (c) altitude; (d) plan curvature; (e) SPI; (f) TWI.







Figure 3. (Continued). (g) TRI; (h) distance from fault; (i) distance from river; (j) land-use/land-cover; (k) rainfall; (l) soil order; and (m) lithology units

Lithology at scale of 1:100000 for the study area was provided by Iranian Department of
Geology Survey (IDGS). Accordingly, 30 classes were used: OMq, PeEf, PlQc, K1bl, Plc, pd,
TRKubl, TRJvm, MPlfgp, OMql, Plbk, E2c, TRKurl, Qft2, MuPlaj, KEpd-gu, Kgu, Qft1, Ekn,
KPeam, PeEtz, Kbgp, EMas-sb, Mgs, TRJlr, Klsol, JKbl, Kur, OMas and Mmn (Figure 3m).
Finally, all the aforementioned groundwater conditioning factors for modeling purposes were
converted to a raster grid with 30 m × 30 m in the ArcGIS 10.2 software.

#### **3.2.** Spatial relationship between spring location and conditioning factors

To assess the spatial relationship between the spring locations and these conditioning factors, in this research, Step-wise Assessment Ratio Analysis (SWARA) (Keršuliene et al., 2010), a Multi-Criteria Decision Making (MCDM) was used. SWARA has received great attention in various fields in the last five years (Alimardani et al., 2013; Hong et al., 2017). The working principal of SWARA was briefly described as follows:

*Phase one*: first, the experts will define the problem-solving criteria. By using the practical
 knowledge of the experts, the priority for the criteria is determined and these criteria are
 organized in descending order.

227 *Phase two*: the following trends is employed for estimating the weight for the criteria:

Starting from the second criterion, the respondent explains the relative importance of the criterion j in relation to the (j-1) criterion, and for each particular criterion as well. As Keršuliene mentioned, this process specifies the Comparative Importance of the Average Value,  $S_j$  as follows (Keršuliene et al., 2010):

$$S_j = \frac{\sum_{i=1}^n A_i}{n} \tag{1}$$

where *n* is the number of experts;  $A_i$  explicates the offered ranks for each factor by the experts; j stands for the number of the factor.

235 Subsequently, the coefficient  $K_i$  is determined as follows:

236 
$$K_j = \begin{cases} 1 & j = 1 \\ S_j + 1 & j > 1 \end{cases}$$
 (2)

237 Recalculation of weight  $Q_i$  is done as follows:

238 
$$Q_j = \frac{X_{j-1}}{K_j}$$
 (3)

239 The relative weights of the evaluation criteria are calculated by the following equation:

240 
$$W_j = \frac{Q_j}{\sum_{j=1}^m Q_j}$$
 (4)

where  $W_i$  shows the relative weight of j-th criterion, and m is the total number of criteria.

# 242 **3.3.** Groundwater spring prediction modeling

As mentioned earlier, in this research, five new metaheuristic optimization algorithms (IWO, DE, FA, PSO, and BA) were investigated for optimizing the parameters of ANFIS. This section briefly presents the theoretical background of these algorithms and ANFIS.

# 246 3.3.1. Adaptive Neuro-Fuzzy Inference System

Adaptive Neuro-Fuzzy Inference System (ANFIS) was proposed by Jang (1993) to solve nonlinear and complex problems in one framework. ANFIS convert input data to fuzzy inputs by using membership function, Also, there are a different membership functions that describing the system behavior (Jang, 1993). ANFIS applies to the Takagi-Sugeno-Kang (TSK) fuzzy model with two "If-Then" rules both having two inputs  $x_1$  and  $x_2$ , and one output f (Takagi and Sugeno, 1985), as follows:

253 Rule 1: if  $x_2$  is  $A_2$  and  $x_2$  is  $B_2$ , then  $f_2 = p_2 x_2 + q_2 x_2 + r_2$  (5)

254 Rule 2 1: if 
$$x_1$$
 is  $A_1$  and  $x_2$  is  $B_1$ , then  $f_1 = p_1 x_1 + q_1 x_2 + r_1$  (6)

Jang's ANFIS consists of feed-forward neural network with six distinct layers. Detailed
 description of ANFIS can be found in (Jangs, 1993).

# 257 **3.3.2.** Meta-heuristic optimization algorithms

The main goal of these algorithms is to find the optimal antecedent and the consequent parameters of the ANFIS model using IWO, DE, FA, PSO, and BA algorithms. Figure 4 illustrates a general methodological data flow of the ANFIS model.





Figure 4. General methodological flow of ANFIS.

# 263 3.3.2.1. Invasive Weed Optimization algorithm

Invasive weed optimization (IWO) mimics the colonizing behavior of weeds. Its design is based on the way to find proper place for growth and reproduction of weeds by Mehrabian and Lucas (2006). One characteristic of this algorithm is its simple structure; the number of input parameters is low and it has strong robustness. Furthermore, it is easy to understand and the same merit causes it to be used for solving difficult nonlinear optimization problems (Ghasemi et al., 2014; Naidu and Ojha, 2015; Zhou et al., 2015). This algorithm consists of 4 parts: initialization, reproduction, spatial dispersal, competitive exclusion and termination condition.

# 271 **3.3.2.2. Differential Evolution algorithm**

272 Differential Evolution (DE) is an evolutionary algorithm for finding global optimal answers for

- problems with continuous space (Das et al., 2009). This algorithm starts by producing a random
- 274 population in which each individual of the population is a solution to the problem. Vector  $X_i^G =$
- 275  $(x_{1,i}^G, x_{2,i}^G, x_{3,i}^G, ..., x_{D,i}^G)$  shows each individual of the population,  $i = \{0, 1, 2, ..., NP\}$  is a number
- denoting each individual, in which D stands for the search dimension, or in other words, is a component problem and  $G = \{0, 1, 2, ..., G_{max}\}$  generation time that  $G_{max}$  is the total number of
- 278 generations.
- By assuming the maximum and minimum of every dimension of searching space, there are  $X_L =$
- 280  $(x_{1,L}, x_{2,L}, ..., x_{D,L})$  and  $X_U = (x_{2,U}, ..., x_{D,U})$ , respectively; initial population is defined as the
- following (Storn and Price, 1997):

282 
$$x_{j,i}^0 = x_{j,L} + rand(0,1).(x_{j,U} - x_{j,L})$$
 (7)

where rand(0,1) is a uniformly distributed random number in [0, 1]. Detailed description of DE can be found in (Chen et al., 2017a).

# 285 **3.3.2.3.** Firefly Algorithm

286 Firefly Algorithm (FA) is as a meta-heuristic algorithm that is originated from flashing and communication behavior of fireflies proposed by Yang (2010). Like other swarm intelligence 287 288 algorithms, where their components are known as solutions for the problems, in this algorithm, each firefly is a solution and its light intensity is the objective function value. In general, FA 289 algorithm follows three idealized rules as below: (1) All firefly species are unisex, with each of 290 them attracting other fireflies without considering their gender (Amiri et al., 2013); (2) 291 292 Attractiveness of a firefly is related to its light intensity, and thus, from two flashing firefly species, one with lower light intensity moves toward the other one with higher light intensity: (3) 293 294 Light intensity of a firefly is defined as an objective function value and must be optimized.

# **3.3.2.4. Particle Swarm Optimization algorithm**

Particle Swarm Optimization (PSO) algorithm has been inspired by the way birds use their collective intelligence for finding the best way to get food (Kennedy and Eberhart, 1995). Each bird implemented in this algorithm acts as a particle that is in fact a representative of a solution to the problem. These particles find the optimum answers for the problem by searching in "n" dimensional space, whereas "n" is the number of the problem's parameters. For this purpose, particles were scattered randomly in space at the beginning of algorithm execution. Detailed description of PSO can be found in Kennedy (2011)

#### **303 3.3.2.5. Bee algorithm**

Bee algorithm (BA), which was introduced by Pham (Pham et al., 2005), is inspired by 304 305 foraging behavior of bees' colonies in search of food sources located near the hive. In the initial setup, evenly distributed scout bees are scattered randomly in different 306 directions to identify flower patches. After that, scout bees come back to hive and start 307 308 a specific dance called the waggle dance. This dance is for communicating with others in order to share the information of discovered flower patches. The information 309 indicates direction, distance, and nectar quality of the flower patches, and helps the 310 colony to have proper evaluation of all flower patches. After evaluation, scout bees 311 come back to the location of discovered flower patches with other bees, named recruit 312 bees. Dependent on the distance and the amount of nectar, different number of recruit 313 bees is assigned to each flower patch. In other words, those flower patches with better 314 nectar quality dedicate more recruit bees to themselves. Recruit bees then evaluate the 315 quality of flower patches when performing the harvest process, and leave the flower 316 patches having low quality. Conversely, if the flower patch quality is good, it will be 317 announced during the next waggle dance. 318

#### 319 **3.4. Performance assessment of models**

According to Chung and Fabbari (Chung and Fabbri, 2003), without validation, the 320 321 result (achieved maps) of the models do not have any scientific significance. Prediction capability of these five spatial groundwater models must be evaluated using both 322 success-rate and prediction-rate curves (Hong et al., 2015). Success-rate curves show 323 built model is the groundwater 324 how suitable the for potential assessment (Gaprindashvili et al., 2014). Success-rate curves have been constructed 325 using groundwater potential maps and the number of spring locations used in training dataset 326 (Pradhan et al. 2010). Prediction rate curves constructed using testing dataset 327 demonstrate how good the model is and evaluate the prediction power of the models. 328

Therefore, it can be used for model prediction capabilities (Brenning, 2005). The area under the curve (AUC) of success and prediction rate is the base for assessment accuracy of the groundwater potential models quantitatively (Khosravi et al., 2016a; Khosravi et al., 2016b; Pham et al., 2017b). The AUC value varies from 0.5 to 1; the higher the AUC, the better the prediction capability of models.

In addition, Mean Squared Error (MSE) was further used (Tien Bui et al, 2016) as
follows:

337 
$$MSE = \frac{\sum_{i=1}^{n} (O_i - E_i)^2}{N}$$
 (8)

where  $O_i$  and  $E_i$  are observation (target) and prediction (output) values in both training and testing dataset and N is the total samples in the training or the testing dataset.

#### 340 **3.5. Inferential statistics**

#### **341 3.5.1 Friedman test**

Non-parametric statistical procedures such as Freidman test (Friedman, 1937) can be 342 used regardless of statistical assumptions (Derrac et al., 2011) and do not presuppose 343 the data to be normally distributed. The main aim of this test is to find whether there is 344 a significant difference between the performed models or not. In other words, 345 performing multiple comparisons to detect significant differences between 346 the 347 behaviors of two or more models (Beasley and Zumbo, 2003). The null hypothesis (H0) is that there are no differences among the performance of the groundwater 348 potential models. The higher the P-value, the higher the probability that the null 349 hypothesis is not true since if the p-value is less than the significance level ( $\alpha$ =0.05), 350 the null hypothesis will be rejected. 351

#### 352 **3.6.2 Wilcoxon signed-rank test**

Because Freidman test only illustrates whether there is any difference between the models or not, this test does not provide pairwise comparisons among compared models. Therefore, another non-parametric statistical test named Wilcoxon signed-rank test have been applied. To evaluate the significance of differences between the performed groundwater potential models, the P value and Z value have been used.

#### 358 **4. Result and analysis**

#### 359 4.1. Multi-collinearity diagnosis

Result of the multi-collinearity analysis in this study is shown in Table 1. The analysis revealed that as VIF is less than 5 and the tolerance is greater than 0.1 indicating no multi-collinearity problem exists among conditioning factors.

363

364

365

• •		Collinearity Statistics		
No	Groundwater conditioning factor	Tolerance	VIF	
1	Slope degree	0.231	2.401	
2	Slope aspect	0.206	4.270	
3	Altitude	0.801	2.097	
4	Plan curvature	0.513	1.446	
5	SPI	0.410	1.689	
6	TWI	0.541	2.113	
7	TRI	0.328	1.939	
8	Distance from fault	0.408	2.25	
9	Distance from river	0.212	3.126	
11	Land-use/land-cover	0.296	3.891	
12	Rainfall	0.298	1.686	
13	Soil order	0.205	4.039	
10	Geology (Unit)	0.215	4.150	

#### **Table.1** Multi-collinearity analysis for conditioning factors.

#### **368 4.2. Determination of the most important parameters**

The most common method of Information Gain Ratio (IGR) was applied to identification of the most important conditioning factors. Result shows that all thirteen conditioning factors are effective on groundwater occurrences. The land-use/land-cover factor has the most important impact on groundwater (IGR=0.502) followed by lithology (IGR=0.465), rainfall (IGR=0.421), TWI (IGR=0.400), soil (IGR=0.370), TRI (IGR=0.337), slope degree (IGR=0.317), altitude (IGR=0.287), distance to river (IGR=0.139), aspect (IGR=0.066), plan curvature (IGR=0.0548), distance to fault (IGR=0.0482) and SPI (IGR=0.0323).

#### 4.3. Spatial relationship between springs and the conditioning factors by SWARA method

377 The spatial correlation between the groundwater springs and the conditioning factors is shown in Table 2 (in the appendix). Regarding the slope, the class of 0-5.5 degree shows the highest 378 379 probability (0.45) on spring groundwater occurrences. As the slope degree increases, the 380 probability of spring occurrence is reduced. In the case of slope aspect, the east aspect (0.44) has the most impact on spring occurrences. According to calculated results, in terms of altitude, the 381 springs are the most abundant in the altitude of 1703-2068 m (0.6). The SWARA model is high 382 383 in flat areas (0.4), followed by concave (0.38) and convex (0.2). For SPI, the highest SWARA value is found for the classes of 583969-1330153 (0.46). In the case of the TWI, the SWARA 384 values decrease when the TWI reduces. There is an inverse relationship between TRI and 385 SWARA value, and as the TRI increases, the SWARA value reduces. 386

Regarding distance from the fault, distance less than 2000 m has the highest impact on spring 387 occurrences and with increase in the distance (greater than 2000 m), the probability of spring 388 389 occurrence is reduced. Regarding distance to river, it can be seen that the class of 0-200 m has the highest correlation with the spring occurrence (0.46) and there is an inverse relationship 390 between spring occurrence and SWARA values. In the case of land-use, the highest SWARA 391 values are shown for garden areas (0.219), followed by mixture of garden and agriculture (0.17), 392 agricultural areas (0.12), whereas the lowest SWARA is for bare soil and rock (0.00063). The 393 rainfall between 500 and 600 mm has the highest SWARA value (0.61). The inceptisols class has 394

the highest SWARA values (0.5) followed by rock outcrop/entisols (0.39), rock outcrop/inceptisols (0.056), inceptisoils/Vertisoils (0.028), and badlands (0.014). The highest probability belongs to the highly porous and very good water reservoir karstic oligomiocene and cretaceous pure carbonate formation (OMq and K1bl), the young and poorly consolidated highly porous detrital rock units (PeEf and Plq) and the unconsolidated quaternary alluvium (PlQc).

# 400 4.4. Application of ANFIS ensemble models and model's assessment

In the current study, hybrids of ANFIS model and five meta-heuristic algorithms were designed,
 constructed and implemented in MATLAB 8.0 software. These models were built using the
 training dataset. Weights gained by the SWARA method for each conditioning factor was fed as
 the input for training dataset. The result is shown in Figures 5 and 6.

As it can be seen in Figure 5, MSE of the ANFIS-IWO model, the ANFIS-DE model, the 405 ANFIS-FA model, the ANFIS-PSO model, and the ANFIS-BA model using the training dataset 406 407 are 0.066, 0.066, 0.066, 0.049, and 0.09, respectively. This indicates that the ANFIS-PSO model has the highest performance, whereas the ANFIS-BA model presents the lowest one. The 408 prediction performance of the five models using the validation dataset is shown in Figure 6. MSE 409 410 of the ANFIS-IWO model, the ANFIS-DE model, the ANFIS-FA model, the ANFIS-PSO model, and the ANFIS-BA model are 0.060, 0.060, 0.060, 0.045, and 0.09, respectively. 411 Therefore, it could be concluded that the ANFIS-PSO model and ANFIS-BA model have the 412 413 highest and lowest prediction performances, respectively.



Figure 5. MSE and RMSE values of the five models using the training dataset of: (a) ANFISIWO; (c) ANFIS-DE; (e) ANFIS-FA; (g) ANFIS-PSO; and (l) ANFIS-BA. Frequency errors of
the five models using the train dataset: (b) ANFIS-IWO; (d) ANFIS-DE; (f) ANFIS-FA; (h)
ANFIS-PSO; and (j) ANFIS-BA.



422 Figure 6. MSE and RMSE values of the five models using the validation dataset: (a) ANFIS-

423 WO; (c) ANFIS-DE; (e) ANFIS-FA; (g) ANFIS-PSO; and (l) ANFIS-BA. Frequency errors of

424 the five models using the validation dataset: (b) ANFIS-IWO; (d) ANFIS-DE; (f) ANFIS-FA;

425 (h) ANFIS-PSO; and (j) ANFIS-BA.

However, it should be noticed that, in addition to accuracy, the execution speed of the five models was found significance. To measure this, the running time for 1000 iteration was estimated. The result is shown in Figure 7. It could be seen that the running time of the ANFIS-IWO model, the ANFIS-DE model, the ANFIS-FA model, the ANFIS-PSO model, and the ANFIS-BA model was 8036, 547, 22111, 1050, and 6993 seconds, respectively. It can be concluded that the ANFIS-DE model had the lowest running time and the ANFIS-FA model had the maximum time.



433



Figure 7. Processing time used for training the five models.

On the other hand, it is possible to test how each model achieves convergence in learning. Using the cost function values, a convergence graph for all five models was constructed and shown in Figure 8. The results show that cost function values of the ANFIS-DE model and the ANFIS-BA model were stable from 30 and 95 iterations, indicating a rapid convergence of the models. While the ANFIS-PSO model, the ANFIS-IWO model, and the ANFIS-FA model showed a convergence after 650, 650, and 360 iterations, respectively. This indicates a slow convergence.









#### 443 4.5. Generation of groundwater spring potential maps using ANFIS hybrid models

Once the five models were successfully trained and validated, these models were used to compute groundwater spring indices for all the pixels of the study areas. Then, these indices were exported from MATLAB into ArcGIS10.2 software for generating groundwater spring potential maps. Ultimately, the achieved maps were visualized by five classes: very low, low, moderate, high and very high (Figures 9a, 9b, 9c, 9d, and 9e).

Many methods can be used for determining thresholds for the five classes, manual, equal interval, geometric interval, quantile, natural break and standard deviation. Selection of a method depends on the characteristics of the data and the distribution of the groundwater spring indexes in a histogram (Ayalew and Yamagishi, 2005). If the indexes have a positive or negative skewness, the quantile or natural break classification is proper for indexes classification (Akgun, 2012). In this research, the histogram was checked and the results revealed that the quantile method was better than other methods for index classification.











**Figure 9**. Groundwater spring potential map using: (a) the ANFIS-IWO model; (b) the ANFIS-DE model; (c) the ANFIS-FA model; (d) the ANFIS-PSO model; and (e) the ANFIS-BA model.

### 464 **4.6. Validation and comparisons of the groundwater spring potential map**

The prediction ability and reliability of the five achieved maps have been evaluated using both 465 the training and the validating datasets. The results of the success rate revealed that the ANFIS-466 DE model had the highest AUC value (0.883) followed by the ANFIS-IWO model (0.882), the 467 ANFIS-FA model (0.882), the ANFIS-PSO model (0.871), and the ANFIS-BA model (0.852) 468 (Figure 10a). Regarding the prediction rate, all five models had a good prediction capability but 469 470 the ANFIS-DE model has the highest prediction rate (0.873) followed by the NFIS-IWO model (0.873) and the ANFIS-FA model (0.873), the ANFIS-PSO model (0.865), and the ANFIS-BA 471 472 model (0.839), respectively (Figure 10b).





Figure 10. (a) Success rate and (b) prediction rate of the five models.



The result of the Friedman test (Table.3) revealed that as Sig and chi-square values were less than 0.05 and greater than 3.84, respectively, the null hypothesis has been rejected. The result also indicated that there was a statistically significant difference between prediction capabilities of the five models.

No	Performed models	Mean rank	<b>Chi-square</b>	Sig
1	ANFIS-DE	3.04		
2	ANFIS-IWO	3.13	<	0.00
3	ANFIS-FA	2.98	64.84	0.00
4	ANFIS-PSO	2.72		
5	ANFIS-BA	3.12		

481 **Table 3**. Result of Freidman test.

482 The results of the Wilcoxon signed-rank test showed that both P-values and z-values were far

from the standard values of 0.05 and (from -1.96 to +1.96), respectively, except for the ANFIS-

FA model vs. the ANFIS-DE model and the ANFIS-PSO model vs. the ANFIS-DE model (Table
485 4). This indicates that there are statistically significant differences between models performance,

485 4). This indicates that there are statistically significant differences between models performance

486 except for ANFIS-FA vs. ANFIS-DE and ANFIS-PSO vs. ANFIS-DE.

487 **Table 4**. Result of Wilcoxon signed rank test.

No	Pairwise comparison	<b>Z-Value</b>	<b>P-Value</b>	Significance
1	ANFIS-DE vs. ANFIS-BA	-3.97	0.00	Yes
2	ANFIS-FA vs. ANFIS-BA	-2.37	0.017	Yes
3	ANFIS-IWO vs. ANFIS-BA	-2.35	0.018	Yes
4	ANFIS-PSO vs. ANFIS-BA	-3.04	0.002	Yes
5	ANFIS-FA vs. ANFIS-DE	-1.32	0.185	No
6	ANFIS-IWO vs. ANFIS-DE	-3.96	0.00	Yes

7	ANFIS-PSO vs. ANFIS-DE	-0.841	0.41	NO
8	ANFIS-IWO vs. ANFIS-FA	-3.19	0.001	Yes
9	ANFIS-PSO vs. ANFIS-FA	-1.90	0.057	Yes
10	ANFIS-PSO vs. ANFIS-IWO	-2.44	0.015	Yes

#### 488 **4.8. Percentage area**

The percentage area of each class of final map resulting from the five hybrid models is shown in Figure 11. According to results of the ANFIS-DE as a most accurate models in groundwater spring potential mapping, the percentage areas of very low, low, moderate, high and very high groundwater spring potential are about 19.06, 19.88, 21.72, 20.55 and 18.78 % of the study area, respectively.



#### 494

495 Figure 11. Percentage areas of different groundwater spring potential classes for five models

#### 496 **5. Discussion**

#### 497 **5.1.** The impact of conditioning factor classes on GSPM

Assessment of conditioning factor is a necessary step in finding the correlation analysis between 498 the groundwater spring and the conditioning factors. It should be noted that no universal 499 guideline is available regarding the number and size of the classes as well as selecting the 500 conditioning factors. They were selected mostly based on characteristics of the study area and 501 previous similar studies (Xu et al., 2013). As the slope increase, the probability of water 502 infiltration reduces and runoff generation will increase. Thus, the steeper the slope, the lower the 503 spring occurrence probability is. According to the result of the SWARA method, the springs 504 almost occur in a middle altitude or mountain slopes. The flat curvature class retains and 505 infiltrates rainfall. Therefore, the amount of groundwater in these areas is higher than concave or 506 convex curvature. The east aspect has more springs than other aspects. These results are in 507 accordance with Pourtaghi and Pourghasemi (2014) who reported that most springs occurred in 508 509 the elevation of 1600-1900 m and east slope aspect (with FR method).

510 TWI shows the amount of wetness, and it is obvious that the more the TWI, the higher the 511 groundwater springs probability occurrence is. Terrain Roughness Index (TRI) or topographic roughness or terrain ruggedness, calculates the sum of change in elevation between a grid cell and its neighborhood, and as the less the roughness, the higher spring potential mapping. The SPI shows the erosive power of the water and mountainous area is higher than plain area. So, as the SPI increases, the spring potential occurrence increases. Rivers are one of the most important sources of groundwater recharge and the nearer to river, the higher probability to springs occurrences. Also, as the rainfall increases, the higher groundwater springs incident, but in the current study, some other conditioning factors affected the spring occurrences.

Most of the springs were located in the garden land-use/land-cover. Therefore, it can be stated 519 520 that the gardens have been established near the springs. Pliocene-Quaternary formation in a geologic time scale is newer and Quaternary formation has a high potential to groundwater 521 springs incident due to high permeability. The fault is discontinuity in a volume of rock. Thus, 522 the nearer to the fault, the higher the spring occurrence probability will be. Inceptisols soils are 523 relatively new and are characterized by having only the weakest appearance of horizons, the 524 most abundant on the Earth (https://www.britannica.com/science/Inceptisol) and mostly formed 525 from colluvial and alluvial materials. So, due to high permeability and high rainfall infiltration, 526 they have a high potential for springs occurrences. In the case of lithological unit, there are four 527 suitable rock type as water reservoir based on physical phenomena such as porosity and 528 permeability that consist of: 1. unconsolidated sands and gravels; 2. sandstones; 3. lime-stones; 529 and 4. basaltic lava flows. In this study area lithological units include sedimentary rocks mostly 530 carbonate and detrital rocks with cover of alluvium and minor soil. 531

# 532 **5.2.** Advantages/disadvantages of the models and performance analysis

The highest accuracy based on MSE in both the training and validating datasets is for the ANFIS-PSO model. However, based on the AUC of the success rate and the prediction rate, the ANFIS-DE model has the highest performance. The problem with MSE comes from the fact that it is based on the error assessment. But the models should be acted upon holistically based on abilities. AUC is based on the true positive (TP), true negative (TN), false positive (FP) and false negative (FN), and therefore is more accurate than RMSE for comparison (Termeh et al., 2018).

ANFIS has a potential to capture the benefits of both neural network and fuzzy logic in a single framework and can be considered as a robust model. ANFIS had some advantages including: (1) much better learning ability, (2) need for fewer adjustable parameters than those required in other neural network structure, (3) allowing a better integration with other control design methods by its networks and is (4) more flexible (Vahidnia et al., 2010; Isanta Navarro, 2013).

545 Despite several advantages of ANFIS, non-adjutancy of membership function is the biggest 546 disadvantage of this model. Finding the optimal parameter for neural fuzzy model in a 547 membership function is difficult; therefore, the best parameter should be finding other 548 optimization models. This problem was addressed in this paper for being solved by five meta-549 heuristic algorithms, namely IWO, DE, FA, PSO and BA.

550 In the current study, the results showed that DE algorithm optimized the parameter for neural 551 fuzzy model better than the four other algorithms. The main DE algorithm's advantage is its 552 simplicity as it consists of only three parameters called N (size of population), F (mutation 553 parameter) and C (crossover parameter) for controlling the search process (Tvrdik, 2006).

- 554 Advantages of DE algorithm can be explained as follows: (1) Ability to handle nondifferentiable, nonlinear and multimodal cost functions, (2) parallelizability to cope with 555 computation intensive cost functions, (3) good convergence properties, and (4) random sampling 556 557 and combining vectors in the present population for creating vectors for the next generation.
- Finally, it should be noted that each algorithm has some advantages or disadvantages according 558 559 to the optimization problems which can be summarized as:

Some of the advantages of IWO include the way of reproduction, spatial dispersal, and 560 competitive exclusion (Mehrabian and Lucas, 2006) as well as the fact that seeds and their 561 parents are ranked together and those with better fitness survive and become reproductive 562 (Ahmed et al., 2014). This algorithm can benefit from combined advantages of retaining the 563 564 dominant poles and the error minimization (Abu-Al-Nadi et al., 2013).

Bees algorithm doesn't employ any probability approach, but utilizes fitness evaluation to drive 565 the search (Yuce et al., 2013). This algorithm uses a set of parameters that makes it powerful, 566 including the number of scout bees in the selected patches, the number of best patches in the 567 selected patches, the number of elite patches in the selected best patches, the number of recruited 568 bees in the elite patches, the number of recruited bees in the non-elite best patches, the size of 569 neighborhood for each patch, the number of iterations and the difference between the value of 570 first and last iterations. 571

- 572 Firefly Algorithm's (FA) advantages are summarized as: (1) handling highly non-linear, multimodal optimization problems efficiently, (2) not utilizing velocities (3) ability to be integrated 573
- with other optimization techniques as a flexible method, and finally (4) not needing a good initial 574
- solution to beginning of its iteration process. 575

Advantages of Particle Swarm Optimization (PSO) algorithm can be summarized as follows: (1) 576 Particles update themselves with the internal velocity; (2) particles have a memory important to 577 the algorithm, (3) the 'best' particle gives out the information to others, (4) it often produces 578 quality solutions more rapidly than alternative methods, (5) it automatically searches for the 579 optimum solution in the solution space (Wan, 2013). 580

As a result, there isn't any algorithm which works perfectly for all optimization problems, and 581 each algorithm has a different performance accuracy based on different data. New algorithms, 582 therefore, should be applied, tested and finally the most powerful algorithm should be selected; 583 as the conclusion of the research demands. 584

#### **5.3.** Previous works and future work proposal 585

Some research has been carried out in groundwater well or spring potential mapping using 586 bivariate statistical models (Nampak et al., 2014; Guru et al., 2017; Al-Manmi and Rauf, 2016) 587 using random forest (Rahmati et al., 2016) and using boosted regression tree and classification 588 and regression tree (Naghibi et al., 2016). The ANFIS-metaheuristic hybrid models have not 589 been used in groundwater potential mapping. However, these hybrid models have proven 590 efficient in flood susceptibility mapping (Bui et al., 2016; Termeh et al., 2018) and landslide 591 susceptibility mapping (Chen et al., 2017). Bui et al. (Bui et al., 2016) ensemble the ANFIS 592 using two optimization models, namely Genetic (GA) and PSO for the identification of flood 593 prone areas in Vietnam. Razavi Termeh et al. (Termeh et al., 2018), used ANFIS-Ant Colony 594

595 Optimization, ANFIS-GA and ANFIS-PSO in flood susceptibility mapping of Jahrom basin and 596 stated that ANFIS-PSO had higher prediction capabilities than the two other models. Chen et al 597 (2017) applied three hybrid models, namely ANFIS- Genetic Algorithm (GA), ANFIS-598 Differential Evolution (DE) and ANFIS-Particle Swarm Optimization (PSO) for identifying the 599 areas prone to landslides in Hanyuan County, China. The results showed that ANFIS-DE had a 600 higher performance (AUC=0.84) followed by ANFIS-GA (AUC=0.82) and ANFIS-PSO 601 (AUC=0.78).

In general, the results of the present study, as well as previous research, find that by applying hybrid models, better results could be achieved for spatial prediction modeling including groundwater potential mapping. The ensembles of ANFIS by meta-heuristic algorithms can be applied for any spatial prediction modeling such as groundwater potential mapping, flood susceptibility mapping, landslide susceptibility assessment, gully occurrences susceptibility mapping and other endeavors at a regional scale and in other areas.

For future work, it is recommended that (1) the water quality of the Koohdasht-Nourabad plain be investigated and the water quality of areas with high potential be determined for different aspects such as drinking, agricultural and industrial activities, and (2) the groundwater vulnerability assessment should be applied by some common methods including DRASTIC model for which the zones with high potential to groundwater occurrences should be preserved against pollution.

# 614 6. Conclusion

615 Groundwater is the most important natural resource in the world and about 25 percent of all fresh 616 water is estimated as groundwater. Thus, the groundwater potential mapping has been considered 617 as one of the most effective methods for the management of groundwater resources for better 618 exploitation. The main result of the present study can be summarized as:

1) The results showed that although all models had good results, but, the ANFIS-DE had the
highest prediction power (0.875) followed by ANFIS-IWO and ANFIS-FA (0.873), ANFIS-PSO
(0.865) and ANFIS-BA (0.839).

According to the results of the SWARA method, most springs existed in an altitude of 17032068 m, flat curvature, east aspect, TWI of 6.6-7.9, TRI of 0-8.7, SPI of 583969-1330153,
Inceptisols soil, slope of 0-5.5 degree, 0-200 m distance from river, 500-1000 m distance from
fault, rainfall between 500-600 mm, in a garden, in a Pliocene-Quaternary lithological age and
OMq lithology unit at the case study.

627 3) Based on the information gain ratio, the most important factors on the groundwater occurrence
628 are land-use/land-cover, lithology, rainfall and TWI. The least important factors are plan
629 curvature, distance to fault and SPI.

- 4) Based on the ANFIS-DE model, totally 39.33% of the case study have a high and very highgroundwater potential placed at north of the case study.
- The result of the current study is helpful for Iran Water Resources Management Company(IWRMC) for sustainable management of the groundwater resources. Overall, the maps resulting

- 634 from these hybrid artificial intelligence algorithms can be applied for better management of the
- 635 groundwater resources in the study area.
- 636 Acknowledgement
- 637 We would like to thank Dr. Bjørn Kristofersen at University of South-Eastern Norway for
- 638 checking English of the manuscript and also two reviewers and Prof. Dimitri Solomatine, editor
- of the Journal of hydrology and earth system science for their positive and helpful comment.

# 640 **References**

- 641 Abu-Al-Nadi, D. I., Alsmadi, O. M., Abo-Hammour, Z. S., Hawa, M. F., and Rahhal, J. S.: Invasive weed
- optimization for model order reduction of linear MIMO systems, Applied Mathematical Modelling, 37,
   4570-4577, 2013.
- Adiat, K., Nawawi, M., and Abdullah, K.: Assessing the accuracy of GIS-based elementary multi criteria
- 645 decision analysis as a spatial prediction tool-A case of predicting potential zones of sustainable
- 646 groundwater resources, Journal of Hydrology, 440, 75-89, 2012.
- Ahmed, A., Al-Amin, R., and Amin, R.: Design of static synchronous series compensator based damping
   controller employing invasive weed optimization algorithm, SpringerPlus, 3, 394, 2014.
- 649 Akgun, A.: A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria
- decision, and likelihood ratio methods: a case study at İzmir, Turkey, Landslides, 9, 93-106,
- 651 10.1007/s10346-011-0283-7, 2012.
- Al-Manmi, D. A. M., and Rauf, L. F.: Groundwater potential mapping using remote sensing and GIS-
- based, in Halabja City, Kurdistan, Iraq, Arabian Journal of Geosciences, 9, 357, 2016.
- Alimardani, M., Hashemkhani Zolfani, S., Aghdaie, M. H., and Tamošaitienė, J.: A novel hybrid
- 655 SWARA and VIKOR methodology for supplier selection in an agile environment, Technological and
- Economic Development of Economy, 19, 533-548, 2013.
- Amiri, B., Hossain, L., Crawford, J. W., and Wigand, R. T.: Community detection in complex networks:
- Multi–objective enhanced firefly algorithm, Knowledge-Based Systems, 46, 1-11, 2013.
- Ayalew, L., and Yamagishi, H.: The application of GIS-based logistic regression for landslide
- susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan, Geomorphology, 65, 15-31,2005.
- 662 Beasley, T. M., and Zumbo, B. D.: Comparison of aligned Friedman rank and parametric methods for
- testing interactions in split-plot designs, Computational statistics & data analysis, 42, 569-593, 2003.
- Berhanu, B., Seleshi, Y., and Melesse, A. M.: Surface Water and Groundwater Resources of Ethiopia:
- Potentials and Challenges of Water Resources Development, in: Nile River Basin, Springer, 97-117,
   2014.
- Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation, Natural
  Hazards and Earth System Science, 5, 853-862, 2005.
- Bui, D. T., Lofman, O., Revhaug, I., and Dick, O.: Landslide susceptibility analysis in the Hoa Binh
- province of Vietnam using statistical index and logistic regression, Natural hazards, 59, 1413, 2011.
- 671 Bui, D. T., Pradhan, B., Revhaug, I., Nguyen, D. B., Pham, H. V., and Bui, Q. N.: A novel hybrid
- evidential belief function-based fuzzy logic model in spatial prediction of rainfall-induced shallow
- 673 landslides in the Lang Son city area (Vietnam), Geomatics, Natural Hazards and Risk, 6, 243-271, 2015.
- Bui, D. T., Pradhan, B., Nampak, H., Bui, Q.-T., Tran, Q.-A., and Nguyen, Q.-P.: Hybrid artificial
- 675 intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood
- susceptibility modeling in a high-frequency tropical cyclone area using GIS, Journal of Hydrology, 540,
- 677 317-330, 2016.
- 678 Chen, W., Panahi, M., and Pourghasemi, H. R.: Performance evaluation of GIS-based new ensemble data
- 679 mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA),

- 680 differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling,
- 681 CATENA, 157, 310-324, 2017.
- 682 Chung, C.-J. F., and Fabbri, A. G.: Validation of spatial prediction models for landslide hazard mapping,
- 683 Natural Hazards, 30, 451-472, 2003.
- 684 Das, S., Abraham, A., Chakraborty, U. K., and Konar, A.: Differential evolution using a neighborhood-
- based mutation operator, IEEE Transactions on Evolutionary Computation, 13, 526-553, 2009.
- David Keith Todd, and Mays, L. W.: Groundwater Hydrology, 2rd Edition, Wiley, New York, 1980.
- 687 Derrac, J., García, S., Molina, D., and Herrera, F.: A practical tutorial on the use of nonparametric
- statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms, Swarmand Evolutionary Computation, 1, 3-18, 2011.
- 690 Emampholizadeh, S., Moslemi, K., and Karami, G.: Prediction the groundwater level of bastam plain
- 691 (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS), Water
- 692 resources management, 28, 5433-5446, 2014.
- Ercin, A. E., and Hoekstra, A. Y.: Water footprint scenarios for 2050: A global analysis, Environment
- 694 International, 64, 71-82, <u>https://doi.org/10.1016/j.envint.2013.11.019</u>, 2014.
- 695 Fashae, O. A., Tijani, M. N., Talabi, A. O., and Adedeji, O. I.: Delineation of groundwater potential zones
- in the crystalline basement terrain of SW-Nigeria: an integrated GIS and remote sensing approach,
- 697 Applied Water Science, 4, 19-38, 2014.
- 698 Fitts, C. R.: Groundwater science, Academic press, 2002.
- 699 Friedman, M.: The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of
- Variance, Journal of the American Statistical Association, 32, 675-701,
- 701 10.1080/01621459.1937.10503522, 1937.
- Gaprindashvili, G., Guo, J., Daorueang, P., Xin, T., and Rahimy, P.: A new statistic approach towards
- 103 landslide hazard risk assessment, International Journal of Geosciences, 5, 38, 2014.
- Ghasemi, M., Ghavidel, S., Akbari, E., and Vahed, A. A.: Solving non-linear, non-smooth and non-
- convex optimal power flow problems using chaotic invasive weed optimization algorithms based on
- 706 chaos, Energy, 73, 340-353, 2014.
- 707 Guru, B., Seshan, K., and Bera, S.: Frequency ratio model for groundwater potential mapping and its
- sustainable management in cold desert, India, Journal of King Saud University-Science, 29, 333-347,
  2017.
- Hong, H., Pradhan, B., Xu, C., and Bui, D. T.: Spatial prediction of landslide hazard at the Yihuang area
- 711 (China) using two-class kernel logistic regression, alternating decision tree and support vector machines,
   712 Catena, 133, 266-281, 2015.
- Hong, H., Panahi, M., Shirzadi, A., Ma, T., Liu, J., Zhu, A.-X., Chen, W., Kougias, I., and Kazakis, N.:
- Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with
   genetic algorithm and differential evolution, Science of The Total Environment, 2017.
- 716 Israil, M., Al-Hadithi, M., and Singhal, D.: Application of a resistivity survey and geographical
- 717 information system (GIS) analysis for hydrogeological zoning of a piedmont area, Himalayan foothill
- region, India, Hydrogeology journal, 14, 753-759, 2006.
- 719 Jang, J.-S.: ANFIS: adaptive-network-based fuzzy inference system, IEEE transactions on systems, man,
- 720 and cybernetics, 23, 665-685, 1993.
- Jha, M. K., Chowdary, V., and Chowdhury, A.: Groundwater assessment in Salboni Block, West Bengal
- 722 (India) using remote sensing, geographical information system and multi-criteria decision analysis
- techniques, Hydrogeology journal, 18, 1713-1728, 2010.
- Kaliraj, S., Chandrasekar, N., and Magesh, N.: Identification of potential groundwater recharge zones in
- 725 Vaigai upper basin, Tamil Nadu, using GIS-based analytical hierarchical process (AHP) technique,
- 726 Arabian Journal of Geosciences, 7, 1385-1401, 2014.
- 727 Kennedy, J., and Eberhart, R.: Particle swarm optimization, IEEE International of first Conference on
- 728 Neural Networks, in, Perth, Australia, IEEE Press, 1995.
- Kennedy, J.: Particle swarm optimization, in: Encyclopedia of machine learning, Springer, 760-766,
- 730 2011.

- 731 Keršuliene, V., Zavadskas, E. K., and Turskis, Z.: Selection of rational dispute resolution method by
- applying new step-wise weight assessment ratio analysis (SWARA), Journal of business economics and
   management, 11, 243-258, 2010.
- 734 Khosravi, K., Nohani, E., Maroufinia, E., and Pourghasemi, H. R.: A GIS-based flood susceptibility
- assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence
- bivariate statistical models with multi-criteria decision-making technique, Natural Hazards, 83, 947-987,
  2016a.
- 738 Khosravi, K., Pourghasemi, H. R., Chapi, K., and Bahri, M.: Flash flood susceptibility analysis and its
- mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical
- index, and weighting factor models, Environmental monitoring and assessment, 188, 656, 2016b.
- 741 Khosravi, K., Pham, B.T., Chapi. K., Shirzadi, A., et al., 2018. A comparative assessment of decision
- trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Science of the
   total environemnt, 627, 744-755, doi.org/10.1016/j.scitotenv.2018.01.266.
- Li, Y.-F., Xie, M., and Goh, T.-N.: Adaptive ridge regression system for software cost estimating on
- multi-collinear datasets, Journal of Systems and Software, 83, 2332-2343, 2010.
- Lohani, A., Kumar, R., and Singh, R.: Hydrological time series modeling: A comparison between
- adaptive neuro-fuzzy, neural network and autoregressive techniques, Journal of Hydrology, 442, 23-35,
  2012.
- 749 Maiti, S., and Tiwari, R.: A comparative study of artificial neural networks, Bayesian neural networks and
- adaptive neuro-fuzzy inference system in groundwater level prediction, Environmental earth sciences, 71,
- 751 3147-3160, 2014.
- 752 Mehrabian, A. R., and Lucas, C.: A novel numerical optimization algorithm inspired from weed
- colonization, Ecological informatics, 1, 355-366, 2006.
- Mohanty, S., Jha, M. K., Raul, S., Panda, R., and Sudheer, K.: Using artificial neural network approach
- for simultaneous forecasting of weekly groundwater levels at multiple sites, Water Resources
- 756 Management, 29, 5521-5532, 2015.
- 757 Mukherjee, S.: Targeting saline aquifer by remote sensing and geophysical methods in a part of
- 758 Hamirpur-Kanpur, India, Hydrogeol J, 19, 53-64, 1996.
- Naghibi, S. A., Pourghasemi, H. R., Pourtaghi, Z. S., and Rezaei, A.: Groundwater qanat potential
- mapping using frequency ratio and Shannon's entropy models in the Moghan watershed, Iran, EarthScience Informatics, 8, 171-186, 2015.
- 762 Naghibi, S. A., Pourghasemi, H. R., and Dixon, B.: GIS-based groundwater potential mapping using
- boosted regression tree, classification and regression tree, and random forest machine learning models in
  Iran, Environmental monitoring and assessment, 188, 44, 2016.
- 765 Naidu, Y. R., and Ojha, A.: A hybrid version of invasive weed optimization with quadratic
- 766 approximation, Soft Computing, 19, 3581-3598, 2015.
- 767 Nampak, H., Pradhan, B., and Manap, M. A.: Application of GIS based data driven evidential belief
- function model to predict groundwater potential zonation, Journal of Hydrology, 513, 283-300, 2014.
- 769 Nosrati, K., and Van Den Eeckhaut, M.: Assessment of groundwater quality using multivariate statistical
- techniques in Hashtgerd Plain, Iran, Environmental Earth Sciences, 65, 331-344, 2012.
- 771 Nourani, V., Alami, M. T., and Vousoughi, F. D.: Hybrid of SOM-Clustering Method and Wavelet-
- ANFIS Approach to Model and Infill Missing Groundwater Level Data, J. Hydrol. Eng., 21, 05016018,
   2016.
- O'brien, R. M.: A caution regarding rules of thumb for variance inflation factors, Quality & Quantity, 41,
  673-690, 2007.
- Oh, H.-J., Kim, Y.-S., Choi, J.-K., Park, E., and Lee, S.: GIS mapping of regional probabilistic
- groundwater potential in the area of Pohang City, Korea, Journal of Hydrology, 399, 158-172, 2011.
- 778 Ozdemir, A.: Using a binary logistic regression method and GIS for evaluating and mapping the
- groundwater spring potential in the Sultan Mountains (Aksehir, Turkey), Journal of Hydrology, 405, 123-
- 780 136, 2011a.

- 781 Ozdemir, A.: GIS-based groundwater spring potential mapping in the Sultan Mountains (Konya, Turkey)
- using frequency ratio, weights of evidence and logistic regression methods and their comparison, Journal
   of Hydrology, 411, 290-308, 2011b.
- 784 Pham, B. T., Bui, D. T., Pourghasemi, H. R., Indra, P., and Dholakia, M.: Landslide susceptibility
- assessment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of
- naïve bayes, multilayer perceptron neural networks, and functional trees methods, Theoretical and
- 787 Applied Climatology, 128, 255-273, 2017a.
- 788 Pham, B. T., Khosravi, K., and Prakash, I.: Application and comparison of decision tree-based machine
- rearning methods in landside susceptibility assessment at Pauri Garhwal Area, Uttarakhand, India,
- Environmental Processes, 4, 711-730, 2017b.
- Pham, D., Ghanbarzadeh, A., Koc, E., Otri, S., Rahim, S., and Zaidi, M.: The bees algorithm. Technical
  note, Manufacturing Engineering Centre, Cardiff University, UK, 1-57, 2005.
- 793 Pourghasemi, H., Moradi, H., and Aghda, S. F.: Landslide susceptibility mapping by binary logistic
- regression, analytical hierarchy process, and statistical index models and assessment of their
- performances, Natural hazards, 69, 749-779, 2013a.
- 796 Pourghasemi, H. R., Pradhan, B., and Gokceoglu, C.: Application of fuzzy logic and analytical hierarchy
- process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran, Natural hazards, 63, 965-996,
   2012.
- 799 Pourghasemi, H. R., Pradhan, B., Gokceoglu, C., Mohammadi, M., and Moradi, H. R.: Application of
- weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping
   at Haraz watershed, Iran, Arabian Journal of Geosciences, 6, 2351-2365, 2013b.
- 802 Pourghasemi, H. R., and Beheshtirad, M.: Assessment of a data-driven evidential belief function model
- and GIS for groundwater potential mapping in the Koohrang Watershed, Iran, Geocarto International, 30,
   662-685, 2015.
- 805 Pourtaghi, Z. S., and Pourghasemi, H. R.: GIS-based groundwater spring potential assessment and
- mapping in the Birjand Township, southern Khorasan Province, Iran, Hydrogeology Journal, 22, 643-662,
  2014.
- Rahmati, O., Samani, A. N., Mahdavi, M., Pourghasemi, H. R., and Zeinivand, H.: Groundwater potential
  mapping at Kurdistan region of Iran using analytic hierarchy process and GIS, Arabian Journal of
- 810 Geosciences, 8, 7059-7071, 2015.
- 811 Rahmati, O., Pourghasemi, H. R., and Melesse, A. M.: 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, 2016.
- 814 Richey, A. S., Thomas, B. F., Lo, M. H., Reager, J. T., Famiglietti, J. S., Voss, K., Swenson, S., and
- Rodell, M.: Quantifying renewable groundwater stress with GRACE, Water resources research, 51, 52175238, 2015.
- Sander, P., Chesley, M. M., and Minor, T. B.: Groundwater assessment using remote sensing and GIS in a
   rural groundwater project in Ghana: lessons learned, Hydrogeology Journal, 4, 40-49, 1996.
- 819 Siebert, S., Henrich, V., Frenken, K., and Burke, J.: Update of the digital global map of irrigation areas to
- 820 version 5, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany and Food and Agriculture
- 821 Organization of the United Nations, Rome, Italy, 2013.
- 822 Singh, A. K., and Prakash, S. R.: An integrated approach of remote sensing, geophysics and GIS to
- 823 evaluation of groundwater potentiality of Ojhala sub-watershed, Mirjapur district, UP, India, Asian
- 824 conference on GIS, GPS, aerial photography and remote sensing, Bangkok-Thailand, 2002,
- 825 Storn, R., and Price, K.: Differential evolution-a simple and efficient heuristic for global optimization
- 826 over continuous spaces, Journal of global optimization, 11, 341-359, 1997.
- 827 Sun, Y., Wendi, D., Kim, D. E., and Liong, S.-Y.: Application of artificial neural networks in
- groundwater table forecasting–a case study in a Singapore swamp forest, Hydrology and Earth System
  Sciences, 20, 1405-1412, 2016.
- 830 Takagi, T., and Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control,
- 831 IEEE transactions on systems, man, and cybernetics, 116-132, 1985.

- 832 Tehrany, M. S., Pradhan, B., and Jebur, M. N.: Spatial prediction of flood susceptible areas using rule
- based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS,
- 834 Journal of Hydrology, 504, 69-79, 2013.
- 835 Tehrany, M. S., Pradhan, B., and Jebur, M. N.: Flood susceptibility mapping using a novel ensemble
- weights-of-evidence and support vector machine models in GIS, Journal of hydrology, 512, 332-343,
- **837** 2014.
- 838 Termeh, S. V. R., Kornejady, A., Pourghasemi, H. R., and Keesstra, S.: Flood susceptibility mapping
- using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms, Science of
- 840 the Total Environment, 615, 438-451, 2018.
- 841 Tvrdık, J.: Competitive differential evolution and genetic algorithm in GA-DS toolbox, Technical
- 842 Computing Prague, Praha, Humusoft, 99-106, 2006.
- 843 Umar, Z., Pradhan, B., Ahmad, A., Jebur, M. N., and Tehrany, M. S.: Earthquake induced landslide
- susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in
  West Sumatera Province, Indonesia, Catena, 118, 124-135, 2014.
- 846 Waikar, M., and Nilawar, A. P.: Identification of groundwater potential zone using remote sensing and
- GIS technique, Int J Innov Res Sci Eng Technol, 3, 1264-1274, 2014.
- 848 Wan, S.: Entropy-based particle swarm optimization with clustering analysis on landslide susceptibility
- mapping, Environmental earth sciences, 68, 1349-1366, 2013.
- 850 Xu, C., Dai, F., Xu, X., and Lee, Y. H.: GIS-based support vector machine modeling of earthquake-
- triggered landslide susceptibility in the Jianjiang River watershed, China, Geomorphology, 145, 70-80,
  2012.
- Xu, C., Xu, X., Dai, F., Wu, Z., He, H., Shi, F., Wu, X., and Xu, S.: Application of an incomplete
- 854 landslide inventory, logistic regression model and its validation for landslide susceptibility mapping
- related to the May 12, 2008 Wenchuan earthquake of China, Natural hazards, 68, 883-900, 2013.
- 856 Yang, X.-S.: Nature-inspired metaheuristic algorithms, Luniver press, 2010.
- 857 Yuce, B., Packianather, M. S., Mastrocinque, E., Pham, D. T., and Lambiase, A.: Honey bees inspired
- optimization method: the bees algorithm, Insects, 4, 646-662, 2013.
- Zare, M., and Koch, M.: Groundwater level fluctuations simulation and prediction by ANFIS-and hybrid
- 860 Wavelet-ANFIS/Fuzzy C-Means (FCM) clustering models: Application to the Miandarband plain,
- Journal of Hydro-environment Research, 18, 63-76, 2018.
- Zhou, Y., Luo, Q., Chen, H., He, A., and Wu, J.: A discrete invasive weed optimization algorithm for
- solving traveling salesman problem, Neurocomputing, 151, 1227-1236, 2015.
- 864

# 865 Appendix

Table.2. Spatial correlation between conditioning factors and the spring locations by SWARA methods

Factors	Classes	Comparative importance of average value Kj	Coefficient Kj=Sj +1	wj=(X(j-1))/kj	weight wj/ sigma wj
	0 - 5.55		1.000	1.000	0.454
	5.55 - 12.11	0.300	1.300	0.769	0.349
Slope	12.11 - 19.43	1.500	2.500	0.308	0.140
(degree)	19.43 - 28.77	2.000	3.000	0.103	0.047
	28.77 - 64.37	3.500	4.500	0.023	0.010

	East		1.000	1.000	0.448
	North	1.000	2.000	0.500	0.224
Slope aspect	West	0.300	1.300	0.385	0.172
	South	0.100	1.100	0.350	0.156
	Flat	0.8	1.05	0.31	0.121
	1703 - 2068		1.000	1.000	0.608
	1385 - 1703	2.200	3.200	0.313	0.190
Altitude (m)	2068 - 3175	0.800	1.800	0.174	0.106
	531 - 1070	1.000	2.000	0.087	0.053
	1070 - 1385	0.200	1.200	0.072	0.044
	Flat		1.000	1.000	0.408
Plan	concave	0.050	1.050	0.952	0.388
curvature	convex	0.900	1.900	0.501	0.204
	583969.72 - 1330153.27		1.000	1.000	0.466
	227099.33 - 583969.72	1.000	2.000	0.500	0.233
SPI	48664.14 - 227099.33	0.200	1.200	0.417	0.194
	0 - 48664.14	1.000	2.000	0.208	0.097
	1330153.27 - 4136452.25	10.000	11.000	0.019	0.009
	6.64 - 7.92		1.000	1.000	0.471
	5.60 - 6.64	0.700	1.700	0.588	0.277
TWI	7.92 - 11.97	1.300	2.300	0.256	0.120
	4.63 - 5.60	0.100	1.100	0.233	0.110
	2.12 - 4.63	4.000	5.000	0.047	0.022
	0 - 5.59		1.000	1.000	0.544
	5.59 - 12.66	0.800	1.800	0.556	0.302
TRI	12.66 - 20.62	1.500	2.500	0.222	0.121
	20.62 - 30.93	3.000	4.000	0.056	0.030
	30.93 - 75.13	10.000	11.000	0.005	0.003
	0 - 200		1.000	1.000	0.242
Distance from fault	200 - 500	0.050	1.050	0.952	0.231
(m)	500 - 1000	0.100	1.100	0.866	0.210

	1000 - 2000	0.050	1.050	0.825	0.200
	> 2000	0.700	1.700	0.485	0.118
	0 - 200		1.000	1.000	0.464
	200 - 500	1.900	2.900	0.345	0.160
Distance from river	500 - 1000	0.050	1.050	0.328	0.152
(m)	1000 - 2000	0.300	1.300	0.253	0.117
	> 2000	0.100	1.100	0.230	0.107
	Garden		1.000	1.000	0.219
	Mixture of garden and agriculture	0.282	1.282	0.780	0.171
	Agriculture	0.340	1.340	0.582	0.128
	Mixture of poor rangeland and follow	0.419	1.419	0.410	0.090
	Follow	0.233	1.233	0.333	0.073
	Mixture of moderate rangeland and agriculture	0.294	1.294	0.257	0.056
	Mixture of very poor forest	0.124	1.124	0.229	0.050
	Mixture of waterway and vegetation	0.549	1.549	0.148	0.032
	Moderate forest	0.205	1.205	0.122	0.027
Land- use/land-	Mixture of agriculture with dry farming	0.064	1.064	0.115	0.025
cover	Wood-land	0.030	1.030	0.112	0.024
	Good rangeland	0.043	1.043	0.107	0.023
	Rangeland	0.333	1.333	0.080	0.018
	Poor rangeland	0.030	1.030	0.078	0.017
	Poor forest	0.210	1.210	0.065	0.014
	Moderate rangeland	0.281	1.281	0.050	0.011
	Bare soil and rock	0.237	1.237	0.041	0.009
	Dense rangeland	0.278	1.278	0.032	0.007
	Dense-forest	10.000	11.000	0.003	0.001
	Waterway	0.000	1.000	0.003	0.001
	Mixture of agriculture with poor-garden	0.000	1.000	0.003	0.001

	Very poor forest	0.000	1.000	0.003	0.001
	Mixture of moderate forest and agriculture	0.000	1.000	0.003	0.001
	Mixture of low forest and follow,	0.000	1.000	0.003	0.001
	Urban and residential	0.000	1.000	0.003	0.001
	600 - 700		1.000	1.000	0.617
	700 - 800	2.200	3.200	0.313	0.193
Rainfall	800 - 900	0.600	1.600	0.195	0.121
(mm)	500 - 600	1.500	2.500	0.078	0.048
	400 - 500	1.300	2.300	0.034	0.021
	Rock Outcrops/Entisols		1.000	1.000	0.509
	Rock Outcrops/Inceptisols	0.300	1.300	0.769	0.392
Soil order	Inceptisols	5.900	6.900	0.111	0.057
	Inceptisols/Vertisols	1.000	2.000	0.056	0.028
	Bad Lands	1.000	2.000	0.028	0.014
	OMq		1.000	1.000	0.133
	PeEf	0.309	1.309	0.764	0.101
	PlQc	0.253	1.253	0.610	0.081
	K1bl	0.113	1.113	0.548	0.073
	Plc	0.014	1.014	0.541	0.072
	pd	0.059	1.059	0.511	0.068
	TRKubl	0.223	1.223	0.417	0.055
Lithology	TRJvm	0.027	1.027	0.406	0.054
(unit)	MPlfgp	0.048	1.048	0.388	0.051
	OMql	0.015	1.015	0.382	0.051
	Plbk	0.081	1.081	0.353	0.047
	E2c	0.291	1.291	0.274	0.036
	TRKurl	0.059	1.059	0.258	0.034
	Qft2	0.335	1.335	0.194	0.026
	MuPlaj	0.100	1.100	0.176	0.023
	KEpd-gu	0.080	1.080	0.163	0.022

 Kgu	0.566	1.566	0.104	0.014
Qft1	0.064	1.064	0.098	0.013
Ekn	0.109	1.109	0.088	0.012
KPeam	0.027	1.027	0.086	0.011
PeEtz	0.328	1.328	0.065	0.009
Kbgp	0.445	1.445	0.045	0.006
EMas-sb	0.310	1.310	0.034	0.005
Mgs	0.626	1.626	0.021	0.003
TRJlr	10.000	11.000	0.002	0.000
Klsol	0.000	1.000	0.002	0.000
JKbl	0.000	1.000	0.002	0.000
Kur	0.000	1.000	0.002	0.000
OMas	0.000	1.000	0.002	0.000
Mmn	0.000	1.000	0.002	0.000