

# Frequency Ratio Model for Mapping Groundwater Potential Zones Using GIS and Remote Sensing; Medjerda Watershed Tunisia

Fatma Trabelsi, Saro Lee, Slaheddine Khlifi, and Achouak Arfaoui

## Abstract

Groundwater potential mapping and its sustainable development are an important aspect in the Lower Valley of Medjerda (LVM) river sub-basin due to increased water demand for irrigation and domestic use. The main goal of this study is to investigate the application of the probabilistic-based frequency ratio (FR) model in groundwater potential mapping at LVM river in Tunisia using GIS. This study includes the analysis of the spatial relationships between Transmissivity and various hydrological conditioning factors such as elevation, slope, curvature, river, lineament, geology, soil, rainfall, and land use. Eighteen groundwater-related factors were collected and extracted from topographic data, geological data, satellite imagery, and published maps. About 60 groundwater data of transmissivity were randomly split into a training dataset, 70% was used for training the model and the remaining 30% for validation purposes. Finally, the FR coefficients of the hydrological factors were used to generate the groundwater potential map. It was classified into five zones as very high, high, moderate, low, and very low. This information could be used by water decision makers as a guide for groundwater exploration and assessment in the LVM River.

## Keywords

Groundwater potential map • Frequency ratio • GIS • Medjerda • Tunisia

F. Trabelsi (✉) · S. Khlifi · A. Arfaoui  
 RU-Sustainable Management of Water and Soil Resources, Higher School of Engineers of Medjez El Bab (ESIM)\_University of Jendouba, Medjez El Bab, Tunisia  
 e-mail: [trabelsifatma@gmail.com](mailto:trabelsifatma@gmail.com)

S. Lee  
 Geological Research Division, Korea Institutes of Geoscience & Mineral Resources (KIGAM), Daejeon, Republic of Korea

## 1 Introduction

In Tunisia, water resources are already limited in space and time. Their sustainable management is a national priority for ensuring water security and development of the country. The Medjerda basin, subject of this study, is the largest watershed in Tunisia, where groundwater resources are used in conjunction with surface water. While local surface water resources are relatively well managed, groundwater resources are hidden and more difficult to conceptualize, in addition to the lack of groundwater data. Hence, it is necessary to recognize the methods used to approach groundwater management and to predict the groundwater potential at the regional scale. Therefore, groundwater potential mapping (GPM) is essential, and constitutes an effective way to explore these invaluable natural resources [1, 4].

Recently, GIS and Remote Sensing techniques have also provided other cost- and time-effective approaches for the spatial prediction of groundwater productivity [3].

The aim of this study is to evaluate the efficiency of the Frequency Ratio (FR) model for groundwater potential mapping from aquifers of the Lower Valley of Medjerda (LVM) river basin.

## 2 Setting & Methods

### 2.1 Study Area

The study area is the Lower Valley of Medjerda (LVM) sub-basin; its reach extends from the dam of Laaroussia to the outlet of the river in the Mediterranean Sea, with a catchment covering approximately 1656 km<sup>2</sup>. The climate is Mediterranean, characterized by mild, rainy winters and hot, dry summers with no marked prevailing seasons.

In the study area, the sources of water supply are mainly the surface water of the Medjerda River, and to a lesser

extent, groundwater. The groundwater is mainly utilized for irrigation purposes via pumping wells. The most common land use within this plain is irrigated agriculture. From a geological point view, the LVM represents a region of subsidence and collapse whose Mio-Pliocene and Quaternary formations cover the major downstream part of the watershed. Hydrogeologically, the LVM aquifers are mainly lying in alluvial deposits of the river, sometimes in sandstone and calcareous deposits. These aquifers are considered as unconfined in nature and are recharged across their entire surface by rainfall infiltration and Medjerda streams.

## 2.2 Methodology

To carry out the GPM, various data sets were used in the study via satellite images, toposheets, field data, hydrogeological data which were collected from three Regional Commissioner for Agricultural Development (CRDA of Ariana, Bizerte and Manouba), and Rainfall data collected from the National Institute of Meteorology (INM). Sentinel-2A data, from the Copernicus Sentinels Scientific Data Hub, were used to interpret the various groundwater influencing factors such as land use, lineament and drainage. Airborne LIDAR satellite Digital Elevation Model (DEM), with 28 m resolution, has been used for preparing morphometric maps such as slope, TWI, Curvature, aspect, etc. Figure 1 summarizes the adopted methodology.

### 2.2.1 Groundwater Occurrence Characteristics

In order to prepare a groundwater database, hydrodynamic data (Groundwater Level, Transmissivity, Groundwater Specific Capacity: SPC) of 60 pumping tests were collected from drilling wells (CRDA of Ariana, Bizerte and Manouba). Transmissivity is determined based on actual pumping test analysis of groundwater well. Out of these, 42 (70%) groundwater data were randomly selected for training the models and the remaining 18 (30%) were used for the validation purposes.

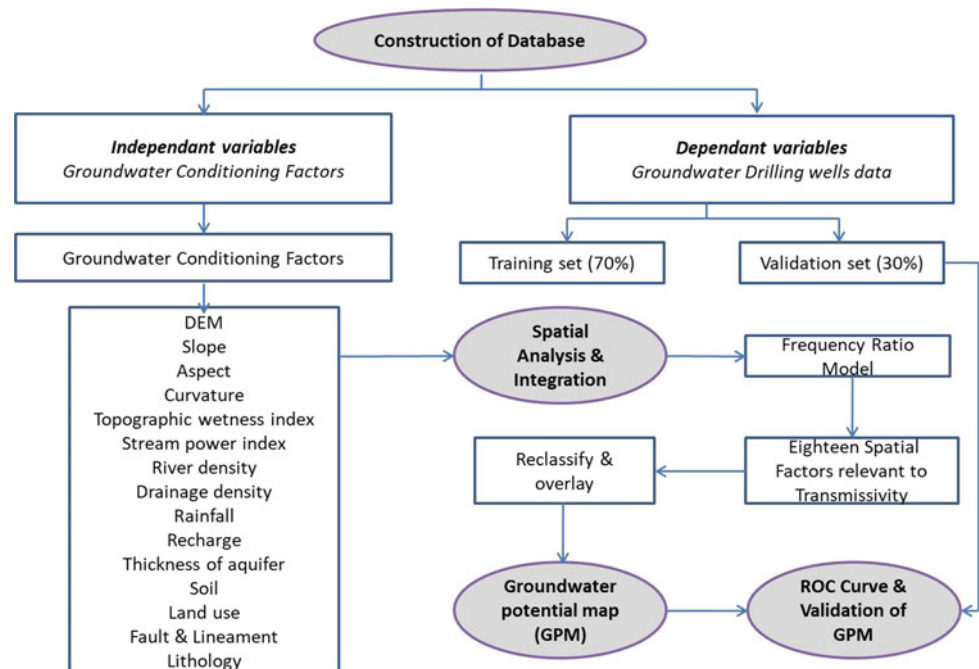
### 2.2.2 Groundwater Conditioning Factors

Generally, the occurrence and productivity of groundwater in a given aquifer is affected by various conditioning factors. The number of conditioning factors used depends on the data availability in the study area. Fifteen thematic layers were chosen as conditioning factors. They provide a reliable data base for an effective groundwater potential prediction of the study area in GIS framework.

### 2.2.3 Frequency Ratio (FR) Model

Frequency ratio (FR) model is a bivariate statistical approach which can be used as a useful geospatial assessment tool to determine the probabilistic relationship between dependent and independent variables, including multi-classified maps [5]. In fact, the FR is defined as the ratio of the area where groundwater wells (high groundwater productivity) occurred in the total study area. FR model structure is based on the

**Fig. 1** Flow chart of methodology



correlation and observed relationships between each groundwater conditioning factor and distribution of groundwater well locations. FR value in each class of the groundwater-related factor can be expressed based on Eq. (1) [6]:

$$FR = \frac{W/G}{M/T} \quad (1)$$

where, **W** is the number of pixels with groundwater well for each conditioning factor; **G** is the number of total groundwater wells in study area; **M** is the number of pixels in the class area of the factor; **T** is the number of total pixels in the study area.

In the FR model, the FR value of each class in a thematic layer was considered as the weight of that particular class in thematic parameters to determine groundwater potentiality.

The FR value was derived from each thematic layer on the basis of a relationship with the Transmissivity values of training data set using Eq. (1). The accuracy of the FR model was validated using receiver operating characteristic (ROC) curve in the form of a success rate curve and a prediction rate curve. The area under the curve (AUC) value indicates the prediction accuracy [4].

In a given pixel, groundwater potential can be determined by summation of pixel values according to Eq. (2) [6]:

$$GWPI = \sum_{R=1}^{R=N} FR \quad (2)$$

where, GWPI and FR are groundwater potential index and the final weight for the FR model, respectively.

The groundwater potential zone indicates the groundwater potentiality Index in quantitative form, i.e. high GWPI indicates high groundwater potentiality whereas low GWPI indicates low groundwater potentiality. In this study the GWPI were derived by integration of all thematic data layer using raster calculator tool in Arc GIS 10.4.

### 3 Results and Discussion

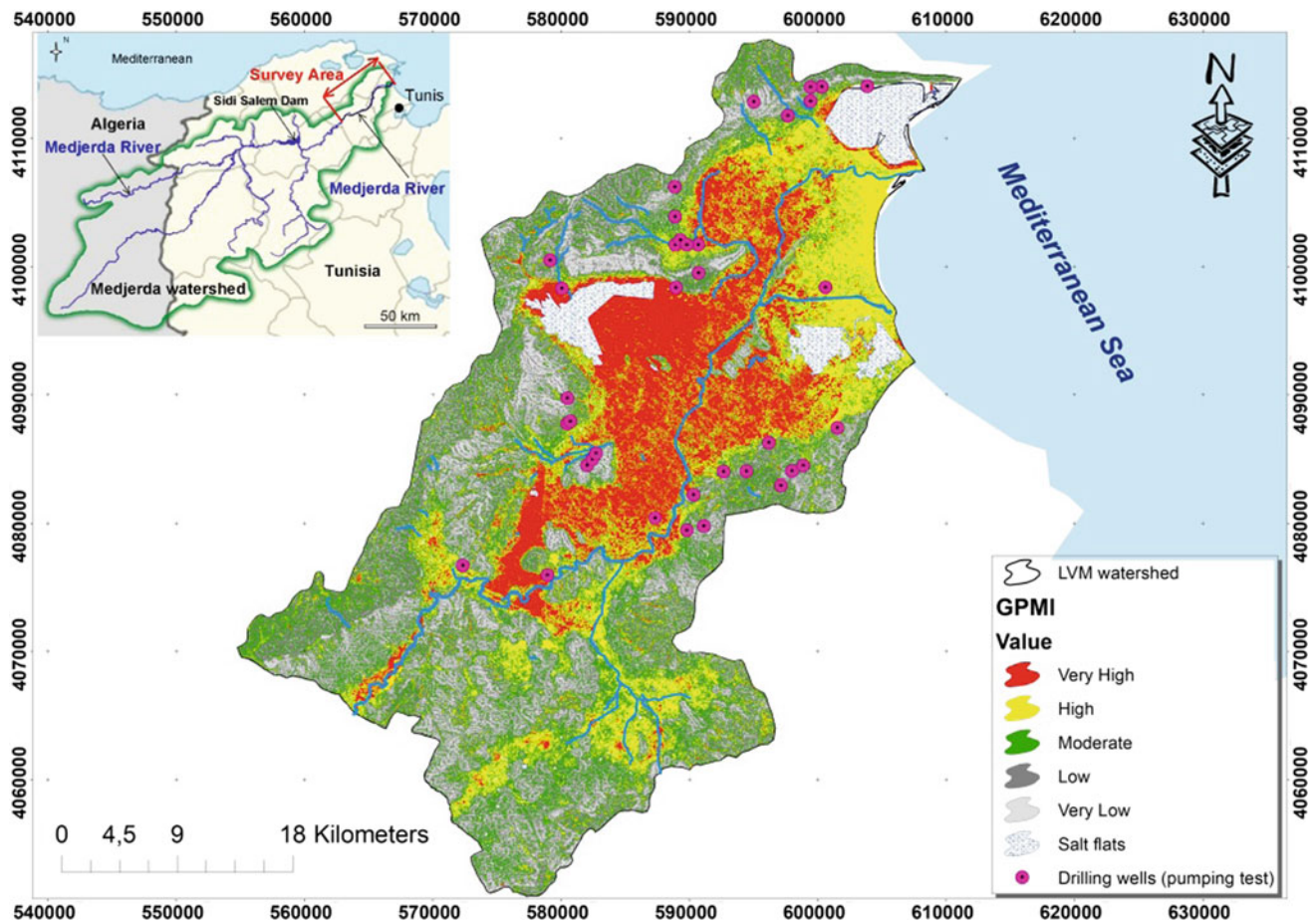
The probability of GW occurrence is usually reduced with the value of altitude. A high slope angle leads to surface runoff increase, and hence the infiltration process will be decreased [6]. In case of slope, the FR ratio is >1 for the first slope range (0°–3°) indicating a high correlation between this slope range and the GW productivity. However, with a high slope range, the ratio suddenly increases along with the slope increase, and then it decreases. To interpret this, it is

important to relate this range with other used factors such as geology. The aerial extension of this range is mainly associated with the extension of flood deposits. These deposits consist mainly of sand and gravel and have higher values of hydraulic conductivity. The higher values of FR for flood deposits support this conclusion. In case of geology, the Quaternary lithological layers have relatively higher values of FR for flood deposits and alluvium. The FRs for the rest of the lithological layers were zero indicating the low probability of groundwater occurrence.

A high slope generates high TWI. There is a good connection between TWI and the occurrence of GW, where the increase in TWI value will result in a higher potential of GW. For the land use, the vegetation indicates a low probability of GW occurrence, while bare soil indicates higher values. The regions that are distant from faults have high FR and vice versa. For rainfall, the high class shows high correlation. Concerning the relationship between groundwater potential and the soil factor, it can be seen that FRs are high for the calcareous soil and low for other groups. The higher infiltration rates of this soil support the resultant higher FR values. As the infiltration rate increases, the groundwater recharge increases as well, leading to more productivity conditions. For the distance-to-river factor, the FR value decreases in relation to the distance to the river.

The final groundwater productivity index for the study area was calculated using Eq. (2) and demonstrated in a map in Fig. 2. The FR model provides values of GWPI ranging from 260 to 3098. The mean value was 68,312 and the standard deviation was 95,337. It was classified into five potential classes based on value intervals given by the natural breaks (Jenks) classification method [2], i.e., very low, low, moderate, high and very high (Fig. 2). It is notable that most of the high and very high potential zones are located in the central part of this study area due to the presence of high lineament density, high river density, lower slope, and porous nature of surface and subsurface which was mainly composed of alluvium, gravel and sandy clay deposits. The high lineament density and low slope categories influence the infiltration of water into the ground in the presence of old and recent alluvium deposits containing sands, clay sands, and gravels in the study area [5, 6].

The low and very low potential zones are found in the surroundings of this central part, where a higher slope and the presence of lower drainage density made the region unfavourable for groundwater potentiality. The integration was taken up both with all thematic data layers and also with its different combinations to evaluate the influence of different causative factors on groundwater potentiality index of the study area.



**Fig. 2** Groundwater potential map

## 4 Conclusions

The FR results showed a high-frequency ratio in areas with slopes  $0^{\circ}$ – $3^{\circ}$ ; lithology of alluvium, sandy clay, sand, and gravel; soil type of calcareous; and high density of river. The high-FR indicates a high groundwater occurrence probability. The validation of a map that used fifteen factors showed an accuracy of 86%. The validation results indicate a viable application of the FR model in groundwater potential mapping. The selection of groundwater factors was based on the availability of the data in the study area. For future research, it is suggested to include many available hydrological factors to improve the prediction accuracy of groundwater potential map. The GPM will be helpful to the decision makers in groundwater management and in identifying suitable locations for drilling production wells.

**Acknowledgments** This research (NRF-2016K1A3A1A09915721) was supported by Science and Technology Internationalization Project through National Research Foundation of Korea (NRF) grant funded by the Korean Ministry of Science and ICT and Tunisian Ministry of Higher Education and Scientific Research.

## References

1. Adiat, K., Nawawi, M., Abdullah, K.: Assessing the accuracy of GIS-based elementary multi criteria decision analysis as a spatial prediction tool—a case of predicting potential zones of sustainable groundwater resources. *J. Hydrol.* **440**, 75–89 (2012)
2. Jenks, G.F.: The data model concept in statistical mapping. *Int. Yearb. Cartography* **7**, 186–190 (1967)
3. Lee, S., Kim, Y.S., Oh, H.J.: Application of a weights-of-evidence method and GIS to regional groundwater productivity potential mapping. *Environ. Manage.* **96**(1), 91–105 (2012)

4. Lee, S., Pradhan, B.: Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* **4**, 33–41 (2007)
5. Manap, M.A., Nampak, H., Pradhan, B., Lee, S., Sulaiman, W.N. A., Ramli, M.F.: Application of probabilistic-based frequency ratio model in groundwater potential mapping using remote sensing data and GIS. *Arab. J. Geosci.* **7**, 711–724 (2014)
6. Oh, H.J., Kim, Y.S., Choi, J.K., Park, E., Lee, S.: GIS mapping of regional probabilistic groundwater potential in the area of Pohang City, Korea. *J. Hydrol.* **399**, 158–172 (2011)