Chapter 11
Artificial Neural Network for Identification of Groundwater Potential Zones in Part of Hugli District, West Bengal, India

Shashank Yadav and Chalantika Laha Salui

Abstract Adequate availability of groundwater for the rural and urban population is essential because groundwater is an important source of drinking water and agricultural irrigation. This study analyzes groundwater potentiality in part of the Hugli District (Goghat-II, Goghat-I, Arambag, Khanakul-I, Khanakul-II, Pursurah, and Tarakeswar blocks) using artificial neural network model (ANN), on GIS platform, and using remote sensing data and some secondary data sources. Factors like elevation, slope, drainage density, flow accumulation, geology, geomorphology, soil, land use land cover, rainfall, pre-monsoon, post-monsoon, and recharge rate were considered to be influencing the groundwater occurrence over this area. Sentinal-2 satellite data, SRTM data processing techniques, and GIS spatial analysis tools were used to prepare these maps. Groundwater depth level data from 34 wells were considered as ground truth data and randomly divided into training and test sets. An ANN based on the relationship between groundwater potential data and the above factors was implemented on R-studio. Each factor's weight and relative importance was determined by the back-propagation training method. Then the groundwater potential indices were calculated, and the final map was created using GIS tools. The resulting groundwater potential map was validated using Area-Under-Curve analysis with data that had not been used for training. An accuracy of 77.78% was obtained. Five categories ("very high," "high," "moderate," "low," "very low") of groundwater potential zones have been demarcated. This groundwater potential information will be useful for effective groundwater management and exploration.

Keywords ANN · Groundwater potential · Sentinel data · SRTM data

S. Yadav · C. L. Salui (✉)
Indian Institute of Engineering Science and Technology, Shibpur, Howrah, West Bengal, India

11.1 Introduction

Groundwater is one of the most valuable natural resources and supports human health, economic development, and ecological diversity. Due to consistent temperature, widespread and continuous availability, excellent natural quality, limited vulnerability, low development cost, and drought reliability, it has become an essential and dependable source of water supplies in all climatic regions (Todd and Mays 2005). But, due to frequent failures in monsoon, undependable surface water, and rapid urbanization and industrialization, considerable risk to this valuable resource was created (Ramamoorthy and Rammohan 2015).Ever-increasing population of the world leads to the increasing demand for surface as well as sub-surface water supply. To meet the demand, there has been indiscriminate exploitation of groundwater resources. Due to the lack of planned groundwater withdrawal approach, random drilling of bore wells results in decrease in groundwater potential, lowering of water level, and deterioration in groundwater quality. Thus it is necessary to develop a sustainable groundwater management system to properly utilize these important resources. The occurrence of groundwater at any place on the earth is a consequence of the interaction of the climatic, geological, hydrological, physiographical, and ecological factors (Acharya and Nag 2013; Antony 2012) and hence is dependent on the correct interpretation of the hydrological indicators and evidence (Biswas et al. 2012). Identification of groundwater potential zones is an essential step to an effective groundwater management. There are several methods for the identification of groundwater potential zones. Remote sensing technique with its advanced features facilitates largely for the purpose. Remote sensing (RS) and geographic information system (GIS) have proven to be efficient, rapid, and cost-effective techniques producing valuable data on geology, geomorphology, lineaments, and slope as well as a systematic integration of these data for exploration and delineation of groundwater potentials zones (Fashae et al. 2014; Kumar et al. 2016). This study was taken up to identify different groundwater potential zones by artificial neural network (ANN) where simulation was based on factors influencing the availability of groundwater along with their test sites. The study was conducted for the part of Hugli district (Fig. 11.1). Understanding the groundwater potential of the area is useful to regulate and optimize water use without adversely impacting water resources for future use in this area. In the Hugli district, the alluvium in the area forms a rich repository of groundwater. Each aquifer system consists of two to three aquifers separated by thin clay layers, which are not regionally extensive (CGWB). The material of the shallow aquifer is fine to medium-grained sands in the upper part and coarser in the lower part. Sand is the dominant lithological unit in the aquifer system of the Hugli district.

Generally, the first aquifer is restricted within 60–80 mbgl (meters below ground level) depth. However, at some places like Tarakeswar (Tarakeswar block) and Bhadreswar (Serampore-Uttarpara block), the first aquifer has been noticed to continue down to 127 mbgl and 113 mbgl depth and further continuing downward, and the second aquifer system starts below 90 mbgl (CGWB). Groundwater occurs
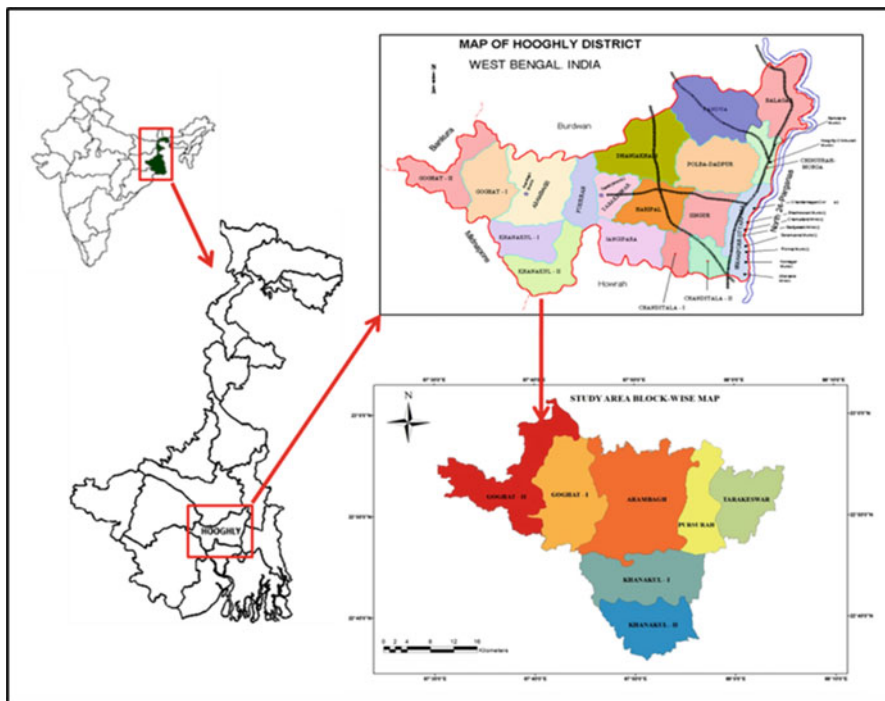
Fig. 11.1 Map showing the location of the study area

in a thick zone of saturation within the alluvium, existing under water-table conditions in the entire area except for a small portion to the west beyond the river Darakeswar.

11.2　Data and Methodology

The data used for this analysis was obtained from the following sources and later modified to meet the requirements of the study using software like ESRI ArcGIS, ERDAS Imagine, QGIS, and R-Studio. For determining groundwater potential zone in the study area, 12 thematic maps, viz., geology, geomorphology, soil, land use/land cover, elevation, slope, drainage density, flow accumulation, recharge, rainfall, and groundwater table depth of pre-monsoon and post-monsoon, were generated using satellite imagery and various conventional datasets (Fig. 11.2). The Digital Elevation Model used in this study is SRTM DEM (Shuttle Radar Topography Mission) and was classified into five elevation classes. Then using Spatial Analyst tool, the slope map was created from this SRTM DEM. For the analysis, it was further sliced into five classes. For the drainage density data, drainages were taken from SOI topographical sheets and were updated from satellite
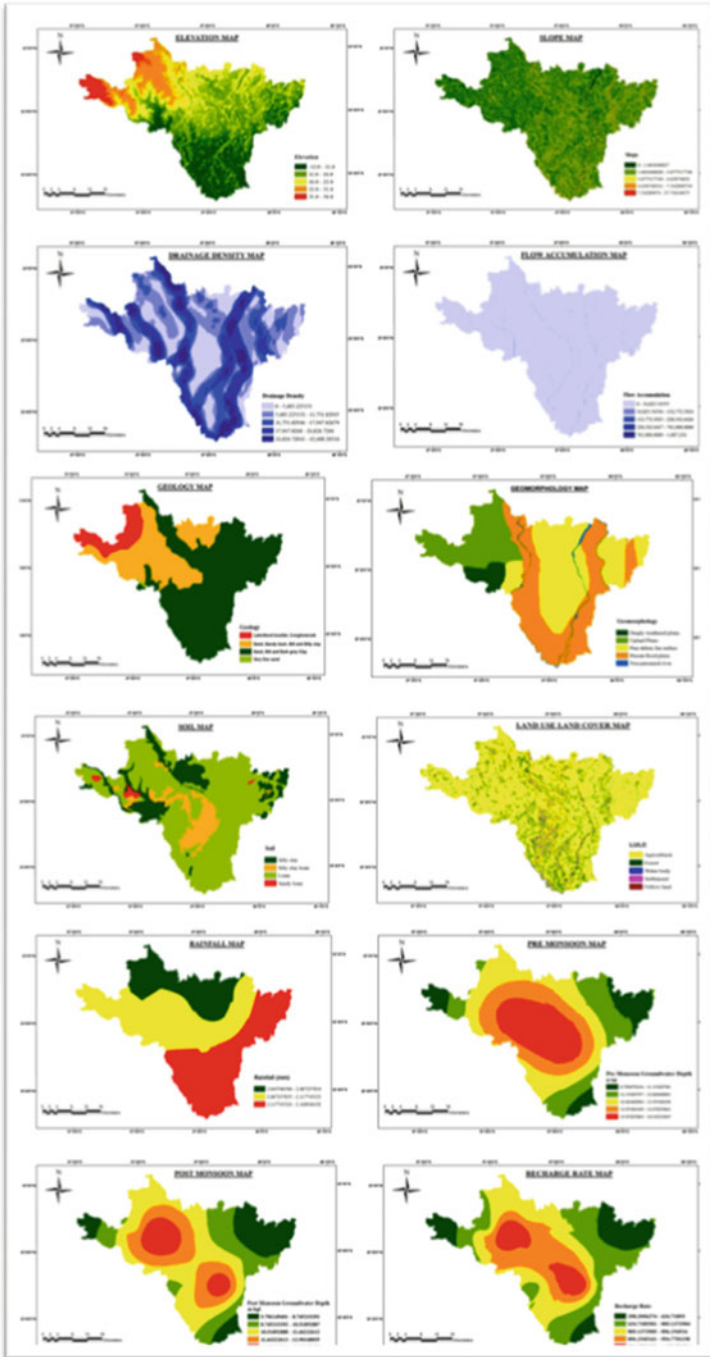
Fig. 11.2 Factors influencing groundwater potentiality

imagery. Then the line density was calculated for each cell. From the combined effect of slope and drainage, flow direction mapping was done followed by the flow accumulation mapping using Arc-hydro tool. The geology information was acquired from the Geological Survey of India; scale of map is 1:250,000. Laterized boulder/conglomerate, sand/sandy loam, silt and silty clay, sand/silt/dark gray clay, and very fine sand are found in this area. Geomorphological units present here are deeply weathered plain, upland plains, para-deltaic fan surface, present flood plains, and non-perennial rivers. The soil map was acquired from the National Bureau of Soil Survey and Land Use Planning (NBSS&LUP) at a scale of 1:500,000. Soil categories like silty clay, silty clay loam, loam, and sandy loam are found in this area. Processing of satellite data was done, so as to make data free from all the errors caused by atmosphere, geometry, and radiometric distortions during the acquisition of data.

Groundwater observation well (especially, depth to the water table) data on spatio-temporal fluctuations of groundwater (with respect to pre and post-monsoon depth) in the study area and its adjoining parts are collected from Central Ground Water Board (CGWB). The annual average pre- and post-monsoon groundwater depth maps were generated by geo-statistical interpolation technique. In this case, the interpolation method was selected to produce pre- and post-monsoon groundwater depth maps and has been extensively used to produce groundwater depth maps.

Annually estimating groundwater storage is based on the water level fluctuation and specific yield approach:

$$\Delta S = \Delta h \times A \times S_y \qquad (11.1)$$

where ΔS=change in groundwater storage, Δh=change in water table elevation during a given period of time, and s_y=specific yield.

In this computational formula, the difference between the highest and the lowest level is taken as the fluctuation of the groundwater level is used for estimating the change in groundwater storage. The area in Eq. (11.1) is the area of influence of the respective monitoring wells in the concerned study area.

11.2.1   Methodology

All of the generated thematic layers were used as independent variables in an artificial neural network for finding the groundwater potential zones (Fig. 11.3). After defining the input and output parameters and then training, validation, and test data set, the learning algorithm was defined, and the network was trained and validated. If results are unsatisfactory, then updating the parameters for the training of the network were followed. After obtaining the best neural network architecture
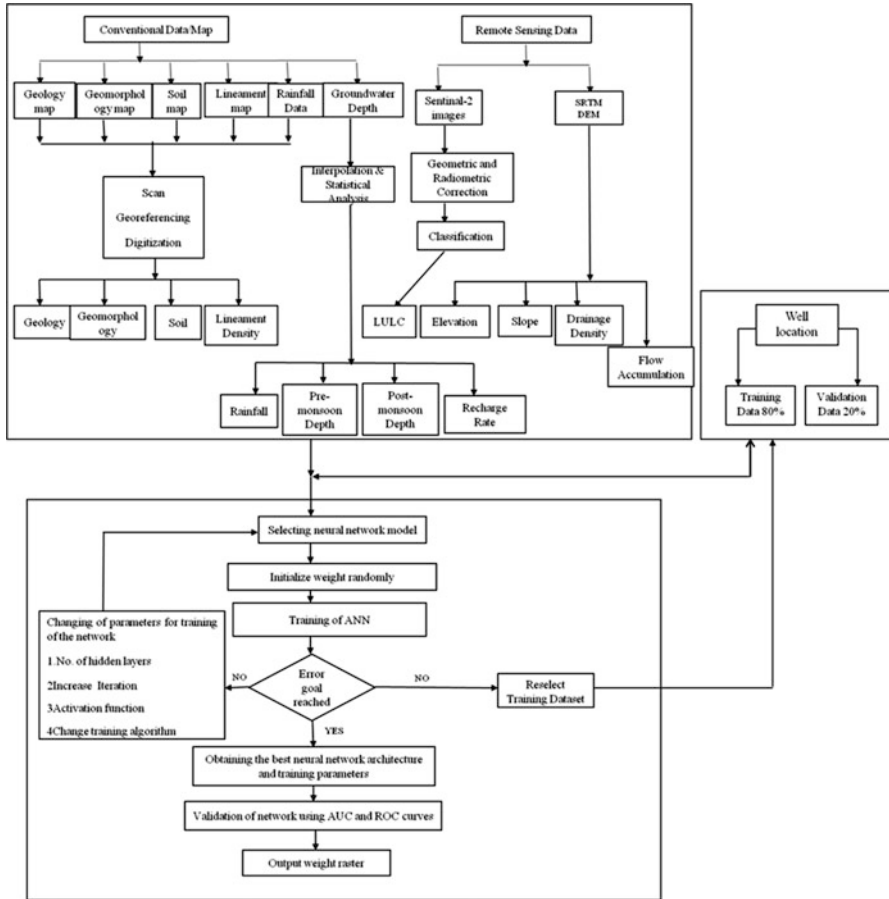
Fig. 11.3 Methodology flowchart

and training parameters with a satisfactory level of accuracy, the final output weight raster was generated and that was classified into various potential zones.

11.3 Results and Discussion

Groundwater well data were collected for pre-monsoon and post-monsoon (34 wells).The main use of the groundwater in this area is agricultural, so our data was obtained between April–June and October–December. Locations considered likely and unlikely to have groundwater were selected as training sites. To select the
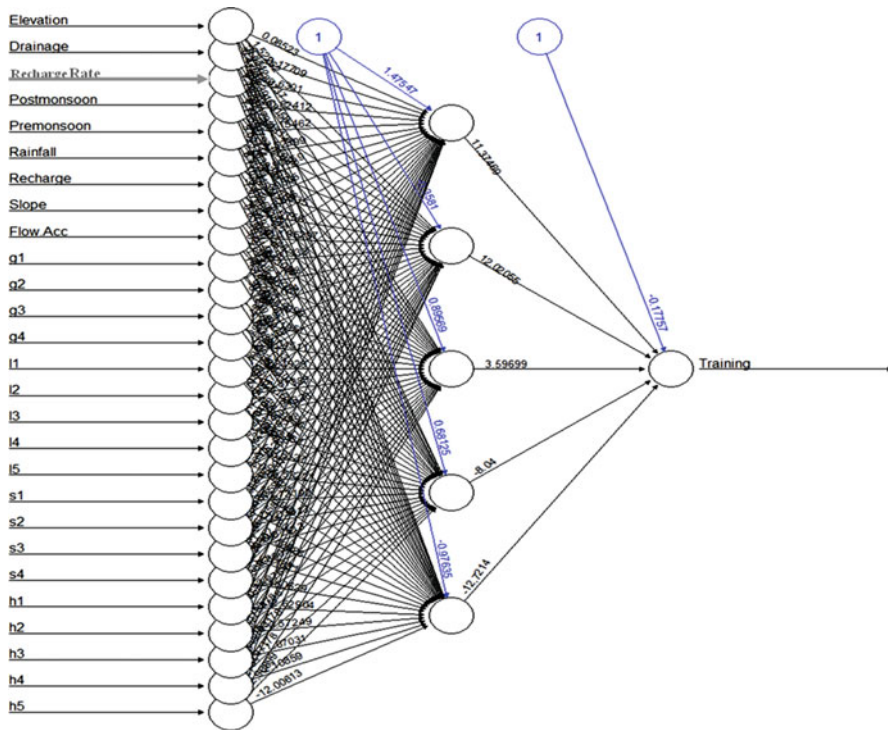
Fig. 11.4 The artificial neural network

training sites based on scientific and objective criteria, pre-monsoon groundwater level was used. 80% of groundwater level values were classified into a groundwater potential training dataset that was randomly selected and used for training. The remaining 20% of groundwater level values were used for validation. csv files of training and testing data were prepared for modelling the Artificial Neural Networking R Studio (Fig. 11.4).

The training and testing data were cleaned up in R Studio. The clean-up process involves removal of empty cells in the data and the conversion of categorical data into numeric data for scaling and computation. The categorical variables of Geology, Geomorphology, Soil, and Land cover were converted to numeric variables of 0s and 1s. The 0s represent the absence of a category in a particular subclass, and the 1s represent the presence of a category in a particular sub-class.

A total of 27 variables were finally available in the training data after conversion of the categorical data into numeric. These variables were then normalized using Z-score standardization or Min-Max scaling as typical neural network algorithms require data that are on a 0–1 scale. The same procedure was applied to the testing data involving reading of testing .csv file, conversion of categorical variables to numeric, and scaling.
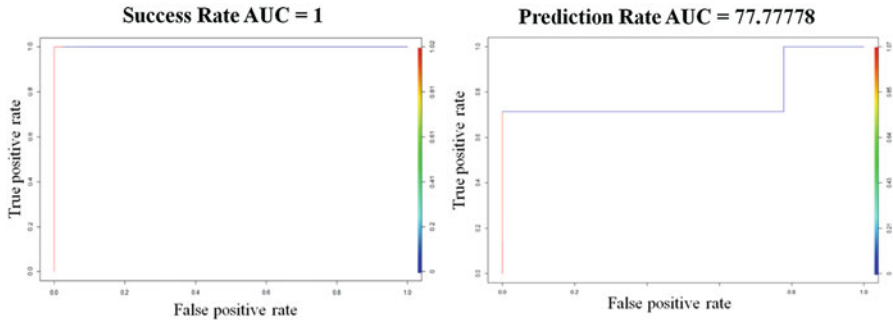
Fig. 11.5 AUC of success and prediction rate ROC

The neural network architecture was set with these parameters:

  (i) No. of hidden layers→ 5
 (ii) Error Function (cost function)→ Cross entropy
(iii) Activation Function→Sigmoid (by default)
 (iv) Algorithm→ Resilient Back Propagation (RPROP+ by default)
  (v) Maximum number of steps →1e+08
 (vi) Threshold→ 0.01 (by default)
(vii) Startweights→ NULL (by default)
(viii) learningrate.limit→ NULL (by default)
 (ix) learningrate.factor→ list(minus = 0.5,plus = 1.2) (by default)

The training error in the artificial neural network was found to be 0.054477 and a threshold of 0.009988.

The trained model was validated with the AUC of the ROC curve. A Receiver Operating Characteristic curve, or ROC curve, is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. The ROC curve is created by plotting the true-positive rate (TPR) against the false-positive rate (FPR) at various threshold settings. The true-positive rate is also known as sensitivity, recall, or probability of detection in machine learning. The false-positive rate is also known as probability of false alarm.

Firstly, the ANN model prediction was checked on the training data itself to see if the model could predict its own data accurately. For this, a code was written in R to compute the output of the ANN model with the training dataset and compare it with the original training data itself, and produce an output table that showed how many 0s were classified as 0s and how many 1s were classified as 1s. The same procedure was then repeated with the testing data to see how many 0s and 1s in the testing dataset were correctly classified and how many were incorrectly classified. After this, the ROCs were plotted, and the AUCs were computed to determine to accuracy of the model (Fig. 11.5).

The success rate represents the accuracy with which the ANN Model predicted the training data itself. This model has a success rate of 1, i.e., 100%. The Prediction
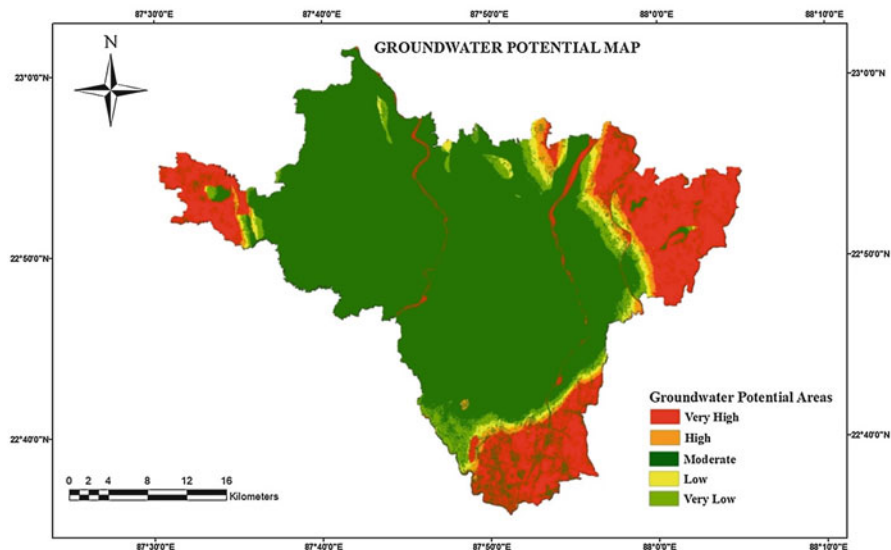
Fig. 11.6 Groundwater potential zone map

rate represents the accuracy with which the ANN Model predicted the testing data. This model has a Prediction rate of 77.78% (Fig. 11.5).

The primary objective of using the groundwater potential map is to effectively predict and display groundwater potential areas. In this study, groundwater potential weight raster was created using the neuron architecture as shown in Fig. 11.6. It is showing three distinct classes representing "good," "moderate," and "poor" groundwater potential zones in the study area. Normally, the "good" groundwater potential zone coincides with high groundwater table which is determined by various factors. River or water bodies are generally considered as the vital sources of recharge of the groundwater table. However, if too many rivers flow in an area, it increases the density of drainage which simultaneously generates high surface runoff and promotes limited infiltration rate. The good groundwater potential zone mainly encompasses Holocene age deposition which is under the present-day flood plain and older flood plain zones around the major river systems. It demarcates the areas where the terrain (flat, gentle slope, soft, unconsolidated) is most suitable for groundwater storage. The "good" groundwater potential zone of the study area covers the north, north-eastern, south, and south-western part, respectively. The "moderate" area of groundwater potential zone is mainly concentrated in the central part and scatter patches randomly distributed over the study area. The hydro-geomorphic feature available in this portion is mainly Holocene to Upper Pleistocene age deposition (western part), a zone of calcareous concentration is present in the upper part of the clay horizon, and the sandy layer often exhibits small sigmoidal, which also suggests moderate capacity of groundwater storage. The "poor" groundwater potential zone
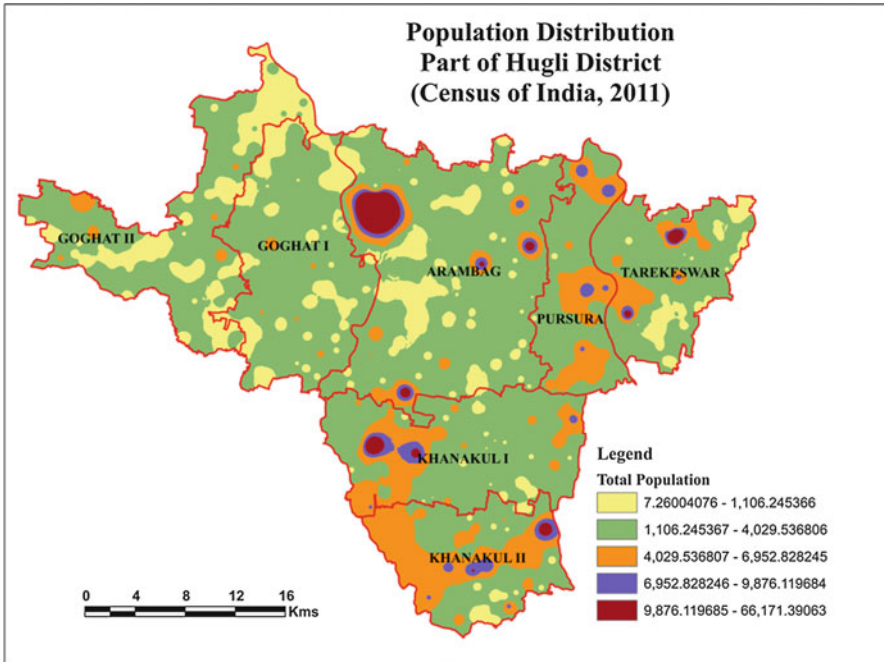
Fig. 11.7 Population density of the study area

has a low discharge (extraction) rate (volume/unit time) as compared to good and moderate groundwater potential zone.

In the study area, the village-wise population was interpolated (Fig. 11.7). High density of population leads to high demand of groundwater. High groundwater potential zones are found in western part of Goghat-II, eastern part of Khanakul-II, and Tarakeswar blocks. High demand areas of Arambag and Khanakul-I are coming under moderate groundwater potential zone. These areas should be taken under groundwater planning and management scheme. Pursurah block should also be monitored as the density of population is moderate for most of the portion of the block with moderate groundwater potentiality.

11.4 Conclusion

In order to sustain long-term agricultural as well as socio-economic development in the Ganga alluvial plain area, judicious use of groundwater is necessary. In this paper, the integrated RS and GIS-based ANN methodology are used to identify groundwater potential zones. By applying these methods, the study area is classified (Hugli district) into five groundwater potential zones, "very high," "high," "moderate," "low," and "very low." The result shows that most parts of the areas with

favorable geology, soil, slope, and optimum rainfall condition have a high potential for groundwater. Identification and selection of a suitable number of thematic layers and justifiable assignment of weights are keys to the benefit of RS and GIS application in determining the potential zone of groundwater resources. Since the methodology adopted in this study is based on logical conditions and it is generic in nature, the same can also be applied in other regions of India or abroad with/without suitable modifications. In addition to designing appropriate policies and institutions toward judicious use and extraction of groundwater resources, there is also a need for a collective approach, particularly by the government organizations, the NGOs, and the common people. The integration of groundwater potential zone, policies, and institutions and the removal of imperfections existing in groundwater use can play a crucial role in this regard. However, a decentralized and participatory planning process is necessary for this purpose.

References

Acharya, T., Nag, S.K., (2013) Study of groundwater prospects of the crystalline rocks in Purulia District, West Bengal, India using remote sensing data. Earth Resource 1(2), 54–59.

Antony, R.A., (2012) Azimuthal square array resistivity method and groundwater exploration in Sanganoor, Coimbatore District, Tamilnadu, India. Res J Recent Sci 1(4), 41–45

Biswas, A., Adarsa, J., Prakash, S.S., (2012) Delineation of Groundwater Potential Zones using Satellite Remote Sensing and Geographic Information System Techniques: A Case study from Ganjam district, Orissa, India Research Journal of Recent Sciences 1(9), 59–66.

Fashae, O.A., Tijani, M.N., Talabi, A.O., Adedeji, OI., (2014) Delineation of groundwater potential zones in the crystalline basement terrain of SW-Nigeria: an integrated GIS and remote sensing approach. App. Wat. Sci.J. 4(1), 19–38.

Kumar, J.R., Dushiyanthan, C., Thiruneelakandan, B., Suresh, R., Vasanth Raja, S., Kumar, S.M., Karthikeyan, K., (2016) Evaluation of Groundwater Potential Zones using Electrical Resistivity Response and Lineament Pattern in Uppodai Sub Basin, Tambaraparani River, Tirunelveli District, Tamilnadu, India. J GeolGeophys 5(2).

Ramamoorthy P., Rammohan V., (2015) Assessment of Groundwater potential zone using remote sensing and GIS in Varahanadhi watershed, Tamilnadu, India, International Journal for Research in Applied Science & Engineering Technology 3(V), 695–702

Todd, D. and Mays, L. (2005) Groundwater Hydrology. 3rd Edition, John Wiley and Sons, Inc., Hoboken.