Fifth Chapter
CHAPTER 5

DATABASE CREATION AND 
GEOSTATISTICAL ANALYSIS USING GIS

5.1 Introduction

Geostatistics is a subset of statistics specialized in analysis and interpretation of geographically referenced data (Goovaerts, 1997). In the most practical perception, geostatistics is a set of techniques/tools used to analyze and predict values of a variable distributed in space and time. One of the main uses of geostatistics is to predict values of a sampled variable over the whole area of interest, referred to as spatial prediction or spatial interpolation. Geostatistics studies spatial variability of regionalized variables. Regionalized variables are those variables which have an attribute value and a location in a two or three dimensional space. The theory of regionalized variables accounts for both the local randomness of variables as well as their spatial structure translated generally as dependence and correlation between different observations.

Geostatistical mapping can be defined as analytical production of maps by using field observations, auxiliary information and a computer program that calculates values at locations of interest. Groundwater levels vary gradually over space and time in a given area. The irregularity of sampling in space and the fact that the groundwater levels exhibit substantial variability with topography, make spatial estimation more difficult. The groundwater modeling needs quantitative information on its behavior in space and time. The large size of the data used in the flow models necessitates the spatial information of groundwater at a cell level in the grid pattern to achieve the reliable results. Spatial analysis of groundwater levels provide more reliable information for the groundwater modeling attempted in the area.
An important distinction between geostatistical mapping and conventional mapping of variables is that the geostatistical prediction is based on application of quantitative, statistical techniques. The traditional approaches to mapping of variables rely on the use of empirical knowledge. In the present time, geostatistics is also increasingly used in combination with GIS to explore spatial variation in remote sensing data, to quantify noise in the images and filling of the missing pixels, to improve generation of DEMs and to optimize spatial sampling.

### 5.2 Point interpolation methods

A point interpolation performs an interpolation on randomly distributed point values and returns regularly distributed point values. The variability of unknown variables is determined by a finite set of inputs and they exactly follow some known physical law. If the algorithm (formula) is known, the values of the target variables can be predicted exactly. Point interpolation assumes also a certain degree of spatial correlation between input point values. Many variables that have discrete values measured at several specific geographic positions can be considered as random processes and can thus be analyzed using spatial autocorrelation analysis, e.g. total rainfall, elevation at triangulated points, grain sizes etc.

There are different approaches to interpolate point data samples, e.g. Deterministic methods, Statistical methods. Methods are deterministic or stochastic, on the basis of how weights are chosen (Hengl, 2007). Stochastic methods use statistical criteria to determine weight factors. Thiessen polygons and inverse distance weighted averaging are the deterministic techniques while polynomial regression, trend surface analysis, and kriging are the stochastic techniques. The deterministic models are flexible and easy to use. They can be considered to be empirical techniques because the user himself selects the parameters of the model, without much deeper statistical analysis. In these types of the models, estimate of the model error is not available and usually no strict assumptions about the variability of a feature exist. Some of the widely used spatial prediction models, such as Thiessen polygons, Inverse distance interpolation, Moving surface and Splines are described below.
Thiessen Polygon method

Thiessen polygon method is one of the simplest techniques of interpolation by drawing boundaries according to the distribution of the sampled data points, with one polygon per data point and the data point located in the center of the polygon. This technique, also referred to as the nearest neighbor method, predicts the attributes of unsampled points based on those of the nearest sampled point and is best for nominal data, where other interpolation methods are not applicable. Each data point represents the part of the area, called the Thiessen polygon, as per its spatial location. All the points falling in the polygon are assigned the variable values equal to that of the data point representing the polygon (Subramanya, 2007). In contrast to this discrete method, all other methods embody a model of continuous spatial change of data, which can be described by a smooth, mathematically delineated surface. It offers a quick way to obtain interpolated point values.

Inverse distance (Moving average) interpolation

The moving average method performs a weighted averaging on point values of a point. The output value for a pixel is calculated as the sum of the products of weights and point values, divided by the sum of weights. Weight values are calculated in such a way that, points close to an output point obtain large weights and points further away obtain small weights. Inverse distance interpolation is an exact and convex interpolation method that fits only the continuous model of spatial variation. For large datasets it can be time-consuming so it is often a good idea to set a threshold distance to speed up the calculations (Hengl, 2007). The points that are further away from an output point than the limiting distance obtain zero weight and thus have no influence on the output value for that point.

Moving surface interpolation

Assuming that the values of target variable at some location are function of coordinates, its values can be determined by finding a function which passes through (or close to) the given set of discrete points. This group of techniques is termed as regression on coordinates.
In the moving surface method, for each output grid node, a polynomial surface is fitted to a larger number of points selected by a moving window (circle). The surface fitting is performed by a least squares fit. Moving surface will completely fail to represent discrete changes in space. This method can be criticized for not relying on empirical knowledge about the variation of a variable. It is probably more advisable to use feature-related geographic predictors such as the distance from a coast line, latitude or longitude and similar, instead of mechanically using the x, y coordinates and their transforms. In that context, regression on coordinates can be considered as the least sophisticated spatial prediction technique. The number of points needs to be larger than the number of parameters (Hengl, 2007).

Splines

Spline is a special type of piecewise polynomial and is preferable to simple polynomial interpolation because more parameters can be defined including the amount of smoothing. The smoothing spline function also assumes that there is a measurement error in the data that needs to be smoothed locally. There are many versions and modifications of spline interpolators. The most widely used techniques are thin-plate splines and regularized spline with tension and smoothing.

The tension parameter controls the distance over which the given points influence the resulting surface, while smoothing parameter controls the vertical deviation of the surface from the points. By using an appropriate combination of tension and smoothing, one can produce a surface which accurately fits the empirical knowledge about the expected variation. Splines have shown to be highly suitable for interpolation of densely sampled heights and climatic variables (Hengl, 2007). However, their biggest criticism is inability to incorporate larger amounts of auxiliary maps to model the deterministic part of variation. In addition, the smoothing and tension parameters need to be set by the user.

In general, empirical prediction models are more primitive than the statistical models and often sub-optimal. However, there are situations where they can perform as good as the statistical models.
5.3 Semi-variogram modeling and Kriging method

Statistical spatial prediction models

In the case of statistical models, coefficients/rules used to obtain outputs are derived in an objective way following the theory of probability. Unlike empirical prediction models, in the case of statistical models, it is necessary to follow several statistical data analysis steps before the maps are generated. This makes the whole mapping process more complicated but it eventually helps to: (a) produce more reliable/objective maps, (b) understand the sources of errors in the data and (c) depict problematic areas/points that need to be revisited (Hengl, 2007). The predictions are accompanied with the estimate of the prediction error. The drawback of this approach is that the input dataset usually need to satisfy strict statistical assumptions. There are at least four groups of statistical models, namely kriging models (plain geostatistics); regression-based models, Bayesian-based model, mixed models (regression-kriging). Kriging based geostatistical analysis of groundwater levels is carried out in this research work.

Semi-variogram modeling

In geostatistics, spatial dependence among the data points is a key factor. It exploits the observation that the values of spatially distributed data are spatially correlated; with the values recorded at the close locations being more highly correlated than the values recorded at widely spaced locations. Covariance and correlation are both measures of the similarity of the head and tail values while semivariance is a measure of the dissimilarity. A plot of semivariance against lag distance is called semivariogram. Semivariogram acts as summary information of the variability of the phenomena at the site. Semivariogram modeling attempts to filter out measurement errors.

A set of observed data points is viewed as a single series or realization of a random function. In order to develop inferential procedure, notion of replicability must be created. This is achieved by the assumption of stationarity, which means that the mean of the random function is constant and covariance between two points depend upon their relative positions. In many cases of hydrological data, the experimental variance increases with the size of the area under consideration. A less stringent hypothesis called
the intrinsic hypothesis has been developed in geostatistics to make estimation possible under such conditions. It assumes stationarity in the first and second moments of the differences between the data points (ASCE, 1990a). The stationarity conditions of the intrinsic hypothesis can be stated as:

\[
\frac{1}{m} \sum_{i=1}^{m} [Z(X_i) - Z(X_{i+h})] = 0 \quad \text{.... [5.1]}
\]

and

\[
\text{Var} [Z(X_i) - Z(X_{i+h})] = \frac{1}{m} \sum_{i=1}^{m} \left[(Z(X_i) - Z(X_{i+h}))^2\right]
\]

\[
= 2\Gamma |h| \quad \text{.... [5.2]}
\]

where, \( Z(X_i) \) is the variable value observed at location \( X_i \),

\( h \) is the incremental distance, called the lag distance,

\( \text{Var} \) is the variance, the expectation of the square of the differences,

\( m \) is the number of pairs of observations separated by the distance vector \( (h) \)

and \( 2\Gamma |h| \) is the variogram at lag distance \( h \)

The experimental data can be analyzed by considering the locations of samples. The computed variance \( \Gamma |h| \) is plotted against lag distance \( h \), called experimental semivariogram. The experimental semivariogram is calculated for discrete values of \( (h) = k.d \), using existing data values, where \( k \) = number of increments and \( d \) = the distance increment. Since observed variable values \( (Z) \) are not separated by exact multipliers of \( d \), the points are grouped according to the closest distance vector \( h \). Then the sample semivariogram is computed as:
\[ I|h| = \frac{1}{2m} \sum_{i=1}^{m} [(Z(X_i) - Z(X_i+h))^2] \]  

where,

\( I|h| \) is mean quadratic increment of \( Z \) between two points separated by the distance \( h \).

A typical semivariogram plot showing its different components i.e. Sill, Range and Nugget, is given in Fig. 5.1. Sill is the semivariance value at which the semivariogram levels off. The lag distance at which the semivariogram reaches the sill value is called Range. The range is one of the most important parameters as it is related with the spatial extent of continuity of the phenomenon. Presumably, autocorrelation is essentially zero beyond the range (Bohling, 2005). Theoretically, the semivariogram value at the origin (zero lag) should be zero. The nugget effect results from high variability at short distances that can be caused by lack of samples, or sampling inaccuracy.

Source: Bohling, 2005

Fig. 5.1 A typical semivariogram and its components
The geostatistical interpolation need to represent the discrete values of the experimental semivariogram by a continuous function called semivariogram model. The reason for this is that the kriging algorithm requires accessing semivariogram values for lag distances other than those used in the empirical semivariogram. More importantly, the semivariogram models used in the kriging process need to obey certain numerical properties in order for the kriging equations to be solvable. Therefore, the model most fitting to the experimental plot from the theoretical semivariogram models is selected. Some of the most frequently used theoretical semivariogram models are Linear, Spherical, Exponential, Gaussian and Power (Bohling, 2005). The spherical model actually reaches the specified sill value, at the specified range. The exponential and Gaussian models approach the sill asymptotically, with range representing the practical range, the distance at which the semivariance reaches 95% of the sill value. The parameters of the semivariogram, i.e. nugget, sill and range are obtained from the fitted semivariogram model for the spatial prediction using Kriging. Fig. 5.2 shows a spherical model fitted to an experimental semivariogram.

![Fitted spherical model](image)

**Fig. 5.2** Fitting a semivariogram model to experimental semivariogram
Kriging

Geostatistical estimators, known as Kriging provide statistically unbiased estimates of surface values from a set of observations at recorded locations, using the estimated spatial and temporal covariance model of the observed data. The most commonly applied forms of Kriging use a semivariogram which is the inverse function of the spatial and temporal covariance. This is a key function of geostatistics and represents the variability of the spatial and temporal patterns of physical phenomena. Kriging can be seen as a sophistication of the inverse distance interpolation. The key problem of inverse distance interpolation is to determine how much importance should be given to each neighbor. There should be a way to estimate the weights in an objective way, so the weights reflect the true spatial autocorrelation structure.

Kriging is an optimal estimation technique which provides the Best Linear Unbiased Estimator (BLUE). It assumes that the sampling domain has a local structure and that knowledge of the local structure can improve the accuracy of the estimated value. The local structure is represented by the Semivariogram model. When it is required to predict the values of a variable \( Z \) at a location \( X_0 \), where no observation is made, the Kriging estimator of \( Z \) at a point \( X_0 \) is expressed as a linear combination (ASCE, 1990a; ASCE, 1990b; Matheron, 1971) as shown in equation 5.4.

\[
Z^*(X_0) = \sum_{i=1}^{n} \left( \lambda_i \ast Z(X_i) \right)
\] ...

[5.4]

where, \( Z^*(X_0) \) = Kriged estimated value of the variable \( Z \) at \( X_0 \),

\( Z(X_i) \) = measured value of the variable \( Z \) at \( X_i \),

\( n \) = number of the observation points

and \( \lambda_i \) is the weight associated with \( Z(X_i) \).

In order to calculate the weights (\( \lambda_i \)), some statistical analysis is carried out using spatial variability of the variable. As Kriging estimator is linear, unbiased and optimal, it must satisfy the following conditions:
(i) Unbiasedness:

\[ \frac{1}{n} \sum_{i=1}^{n} [Z(X_0) - Z^*(X_0)] = 0 \quad \text{.... [5.5]} \]

which gives,

\[ \sum_{i=1}^{m} \lambda_i = 1 \quad \text{.... [5.6]} \]

where,

\( Z(X_0) \) is the true value of the variable \( Z \) at \( X_0 \)

\( Z^*(X_0) \) is the estimate of \( Z(X_0) \).

(ii) Optimality:

\[ \text{Var} [Z(X_0) - Z^*(X_0)] \text{ is minimum} \quad \text{.... [5.7]} \]

Developing equation [5.3] with equations [5.5] and [5.6], a matrix relation with the Kriging weights and Semivariogram is obtained as:

\[
\begin{bmatrix}
0 & \Gamma_{12} & \Gamma_{13} & \cdots & \Gamma_{1n} & 1 \\
\Gamma_{21} & 0 & \Gamma_{23} & \cdots & \Gamma_{2n} & 1 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\Gamma_{n1} & \Gamma_{n2} & \Gamma_{n3} & \cdots & 0 & 1 \\
1 & 1 & 1 & \cdots & 1 & 0
\end{bmatrix}
\begin{bmatrix}
\lambda_0^1 \\
\lambda_0^2 \\
\vdots \\
\lambda_0^n \\
\mu
\end{bmatrix} =
\begin{bmatrix}
\Gamma_{10} \\
\Gamma_{20} \\
\vdots \\
\Gamma_{n0} \\
1
\end{bmatrix} \quad \text{.... [5.8]} \]

where,

\( \Gamma_{ij} = \Gamma(X_i - X_j) \), Diagonal element \( \Gamma_{ii} = \Gamma(X_i - X_i) = 0 \),

and \( \mu \) is the Lagrange parameter to introduce the unbiasedness condition.

The square matrix represents a set of \((n+1)\) simultaneous equation with \((n+1)\) unknowns. The solution provides the weights associated with the
measurement points for Kriged estimate at the point \( X_0 \), yielding the minimum error variance. The solution of matrix representing the relation between Kriging weights and Semivariogram yields minimum error variance for the estimated parameter value at any location within the data domain.

5.4 Geographic Information System

Geographic Information System (GIS) is an information tool that helps us to store, organize and utilize spatial information. GIS is now-a-days indispensable in many different fields of applications to assist decision making process. The application of GIS is found in many areas like water resources management, groundwater modeling, environmental monitoring, disaster assessment, estimation of crop production, transportation system management, military and many more.

5.4.1 Integrated Land and Water Information System

Integrated Land and Water Information System (ILWIS) is a user-friendly and widely distributed GIS with image processing capabilities. It has been developed by the International Institute for Aerospace Survey and Earth Sciences, Enschede, The Netherlands. From July 2007, ILWIS is freely available as open source software under the 52° North initiative.

ILWIS (Version 3.7.2) is used to generate the maps for various input parameters of the groundwater model. It incorporates different geostatistical tools for the data analysis in GIS environment. The spatial autocorrelation of point data is possible using semivariogram analysis in ILWIS. The experimental semivariogram achieved from observed data points is plotted and best fitting theoretical semivariogram models is fixed. Different model parameters of the fitted theoretical model are decided for its further use in statistical interpolation. Different point interpolation techniques are available for the randomly distributed data point values and the interpolation performed in ILWIS returns the regularly distributed point values. The various deterministic interpolation methods available in ILWIS are Nearest point (Thiessen polygon), Moving average, Trend surface
and Moving surface. Furthermore, Kriging is the statistical method of interpolation available in ILWIS that gives the interpolated map as well as the error map with the standard errors of the estimates.

5.5 Database Creation for Study Area

The observation points for different parameters of the groundwater flow in the study area are randomly distributed. The use of the finite difference method based MODFLOW software in the groundwater modeling necessitates the parameter values at regularly distributed nodes of the grid. The point interpolation methods supported by ILWIS are used for this purpose and different thematic data layers are generated.

5.5.1 Geo-referencing of the base map

The study area, Matar branch canal command area, is a part of MRBC command area. The MRBC command area map incorporating canal network, important locations, major drains, railways, major roadways and major water tanks, prepared by Irrigation department, GOG, India is taken as the base map. The scale of the map is 1:126720 and it is imported in the GIS environment. The geo-referencing of the base map is carried out using UTM projection system. In geo-referencing the map, datum is taken as Everest 1975 and ellipsoid as Everest (India, 1956). The MRBC command area map used as the base map for the database creation is shown in Fig. 5.3.
Fig. 5.3 Base map of MBC command area for the GIS based analysis
5.5.2 Generation of Thematic Data Layers

Rainfall data layer

The point map for rain gauge stations in MRBC command area is created by digitization of rain gauge locations. Thiessen polygon map for the rain gauge stations located in MRBC area is generated using the nearest point operation in ILWIS as shown in Fig. 5.4. The rainfall data for MRBC command area is obtained from State Water Data Centre (SWDC), Gandhinagar. The recharge estimation is one of the most vital input parameters for groundwater modeling in the study area. Thiessen polygon map for MRBC area is used to incorporate annual rainfall for different years and the rainfall components of the groundwater recharge for the corresponding years are computed for the Matar branch canal command area. It is found that the study area is covered by the rain gauge stations at Matar and Nadiad.

Fig. 5.4 Thiessen polygon map for rain gauge stations in MRBC command area
Data layers for aquifer properties

Gujarat Water Resources Development Corporation (GWRDC) had estimated aquifer parameters in MRBC command area by pump tests at various locations (Appendix – II). The point map for pump test locations is prepared and attributes for aquifer permeability values are assigned to these points. The thematic layer for aquifer permeability is generated using moving average operation in ILWIS. The resulting thematic map represents the aquifer permeability values at regularly spaced points in the area. The thematic data layer generated for aquifer permeability in MRBC command area is used to crop the information for the study area as shown in Fig. 5.5. The thematic data layers for aquifer transmissivity and specific yield are also generated for MRBC area and the maps for the Matar area are extracted. Fig. 5.6 and Fig. 5.7 represent the data layers for transmissivity and specific yield respectively. The available aquifer parameter values at regularly spaced points in the area are the essential input for groundwater flow model.
Fig. 5.5 Hydraulic conductivity (m/day) data layer for Matar canal command area
Fig. 5.6 Aquifer Transmissivity data layer (m\(^2\)/day) for Matar canal command
Fig. 5.7 Aquifer Specific Yield (%) data layer for Matar canal command.
5.6 Geostatistical Analysis for Study Area

Geostatistical techniques have been applied to study variability of physical, chemical and hydrologic soil properties including mapping and modeling of groundwater. GIS assisted geostatistical analysis is carried out to generate different data layers for input parameters in groundwater flow model for the study area. A network of large number monitoring wells for the groundwater depths has been established by Gujarat Engineering Research Institute (GERI), Vadodara in MRBC command area. The pre-monsoon condition of the groundwater table is monitored by the observations of water depths in the wells in the month of May while the post-monsoon condition of the groundwater table is examined by the water depths in the well observed in the month of October every year. Keeping in view of the continuous data record availability and spatial location covering the major part of the area, 137 groundwater monitoring wells are selected for the use of their depth to water levels. Geostatistical analysis of the groundwater depths obtained at these 137 observation points is used to generate the thematic layers for the groundwater levels in MRBC command area.

5.6.1 Groundwater level data layers

The monitoring well locations are digitized from the map and a point map is created in GIS environment using ILWIS. The ground levels (amsl) and pre-monsoon depth to water levels during different years are attributed to the corresponding monitoring well locations in the map. Geostatistics based Kriging method of interpolation is used in ILWIS to generate the groundwater level data layers in the area. Semivariogram modeling is very necessary to decide the parameters of the theoretical semivariogram model, which are the much needed data inputs for Kriging. Groundwater level at a monitoring well is obtained by subtracting the well water depth from the ground level (amsl) of the corresponding well. The semivariogram analysis is performed for the data sets of groundwater levels at 137 monitoring well locations. The experimental semivariogram is plotted from the data sets and different theoretical semivariogram
models are tested for fitting to the experimental plot. It is found that the exponential model \( y = 8.108 e^{0.000083x} \) fits the most to the semivariogram. The model parameters in terms of nugget, sill and range are decided from the plot. The fitted exponential model to the experimental semivariogram for pre-monsoon condition in year 2000 is presented in Fig. 5.8.

**Fig. 5.8** Exponential model fitting to Pre-monsoon groundwater levels in the year 2000
The parameters of the exponential model fitted to the data set for pre-monsoon condition in the year 2000 are used in Kriging operation in ILWIS and the thematic layer for the groundwater levels is generated. The generated thematic layer is accompanied by the estimate of the prediction error in form of the separate map/layer. The groundwater level data layer provides groundwater heads at the regularly spaced points which are used as input cell values in the grid of the groundwater model for the study area. The cropped map of groundwater head surface for Matar branch canal area is shown in Fig. 5.9. The thematic layers for pre-monsoon groundwater levels for all the years during 2001-2010 are also generated for their use in groundwater modeling.

A separate layer of the prediction error is simultaneously generated in ILWIS for each of the thematic layer of the groundwater levels in Kriging to provide the level of accuracy of the interpolation operation. A Kriged error map for the groundwater head layer for pre-monsoon condition in the year 2000 is presented in Fig. 5.10.

The ground levels of the monitoring well locations are used in ILWIS and semi-variogram analysis is performed to decide the best fitting model to the experimental semi-variogram. The exponential model is fitting to the experimental semi-variogram and the model parameters are referred to for its use in Kriging. The thematic layer for aquifer top surface is generated using Kriging as shown in Fig. 5.11.
Fig. 5.9 Pre-monsoon groundwater head surface (m, amsl) for Matar area in the year-2000
Fig. 5.10 Error map for pre-monsoon groundwater head (m, amsl) surface in the year 2000
Fig. 5.11 Thematic layer of aquifer top surface (m, amsl) for Matar area.
5.7 Results and Discussion

GIS assisted geostatistical analysis is carried out to generate thematic data layers for different input parameters of the groundwater model for the study area. Thiessen polygon map for locations of the rain gauge stations in MRBC area is constructed (Fig. 5.4) in ILWIS using nearest point operation and weighing area for each rain gauge station is decided. It is found that the study area falls within the Thiessen polygons for Matar and Nadiad rain gauge stations. The Matar branch canal command area map is superimposed on the Thiessen polygon map of MRBC area. The area falling within the Thiessen polygon of Matar rain gauge station is assigned the rainfall at Matar during the respective years and 17% of the rainfall is taken as rainfall recharge. Similarly, the rainfall recharge for the Matar canal command area falling within the Thiessen polygon of Nadiad rain gauge station is computed with rainfall recorded at Nadiad rain gauge station.

The aquifer properties such as hydraulic conductivity, transmissivity and specific yield are the primary input parameters for the groundwater model. The thematic data layers for aquifer properties are generated using the results of pump tests conducted by GWRDC, Ahmedabad at various locations in MRBC area. These thematic data layers are used to crop the maps for aquifer properties in Matar branch canal command area (Fig. 5.5, 5.6 and 5.7). It is observed that the hydraulic conductivity in the study area varies between 11-49 m/day while transmissivity ranges between 101-320 m²/day. The specific yield of the unconfined aquifer ranges between 8-10%.

The spatio-temporal variability of the groundwater levels is very well addressed by the geostatistical analysis using Kriging. The semivariogram modeling performed to decide the theoretical model fitting the experimental semivariogram and it is found that the exponential model \( y = 8.108 e^{0.000083x} \) fits the most. The derived parameters of the exponential model are used in Kriging. Various thematic layers of the groundwater head for pre-monsoon condition during the years 2000-2010 are also generated (Fig. 5.9).
The prediction error maps are also generated together with the groundwater level surfaces. These maps shows amount of the error at different locations in the area. Referring to the prediction error maps (Fig. 5.10), it is observed that the error involved in the interpolation is reasonably acceptable for the use of the results in groundwater modeling. The groundwater head layers are used to derive node values in the finite difference grid for the groundwater model. Further, the thematic layer for aquifer top surface is generated using the ground levels of the monitoring wells for its use in the groundwater model (Fig. 5.11).

5.8 Summary and Conclusions

The natural resource inventories need to be regularly updated or improved in detail. It is often needed to consider collection of new samples or additional samples that are then used to update an existing GIS layer. A large data set of groundwater levels exhibiting spatial and temporal variations poses a big challenge for the researchers working with the groundwater flow models. GIS assisted geo-statistical data analysis proves to be an important exercise for this purpose.