Statistical Analysis of Inter-correlation between Pumping Well Data

A Class Project (Geol659)
Submitted
To
Joe Donovan

By
Madan Maharjan
Department of Geology and Geography
West Virginia University
2012
Abstract

Aquifer parameters, transmissivity (T) and storativity (S), define the nature of an aquifer. It is a primary task to determine these parameters before developing a program for sustainable management of an aquifer. Pumping test is often performed to estimate hydraulic properties of an aquifer system like transmissivity, hydraulic conductivity, and storativity. The main purposes of this class project were to statistically analyze the dependent and independent variables of pumping well test data using principal component analysis (PCA) and cluster analysis and later depict the clusters spatially. The result of the analysis demonstrated that the thickness of an aquifer had the most significant impact on the aquifer properties especially due to fractured bedrock and thick overlying glacial outwash. Water table depth and potentiometric surface lower as glacial thickness and well depth increased so as to produce lower hydraulic conductivity and higher drawdown. PCA and cluster analysis enabled to distinguish the least and the most productive wells sites which were supported by presence of maximum numbers of wells within the group. Furthermore, it was successfully recognized Auburn Township as the most productive Township based on the frequency distribution of productive wells within the study area.

Introduction

Aquifer parameters, transmissivity (T) and storativity (S), define the nature of an aquifer. It is a primary task to determine these parameters before developing a program for sustainable management of an aquifer. However, prediction of spatial distribution of hydraulic properties within heterogeneous sedimentary bedrock, especially when fractured, is never easy and remains a challenge (National Research Council 1996). Pumping test is one of key tests done to estimate hydraulic properties of an aquifer system like transmissivity, hydraulic conductivity, and storativity. It is a controlled field experiment in which a well is pumped at a constant rate and drawdown is measured in one or more surrounding observation wells and optionally in the pumped well itself.

The data for this class project was taken from my Master's Degree research. The research was focused on mapping hydraulic conductivity to test the feasibility of locating transmissive bedrock fractured zones under the blanket of glacial till and define their spatial trends (if any) using available pumping test data from Well Logs and Drilling Reports of private residential water wells acquired from Ohio Department of Natural Resources (ODNR), and further analyze the origin of the fractures in relation to tectonic history of the study area.

The study results showed that identifying hydraulic conductivity in fractured rock formations was closely related to fracture characteristics where fractures of all sizes have a dominant effect on the transport properties in such tight sedimentary rock formations. Although, there were local spatial variability in hydraulic conductivities, higher values of hydraulic conductivity were depicted in two distinct orientations (N34°E and N44°W) representing the regional fracture pattern (Maharjan, 2011).

For this class project, the data were used in different perspective so as to analyze statistically and to find inter- correlation between two groups of data. Furthermore, the spatial distribution of the data was done to study the productivity of the wells. The two groups of data were: 1) dependent and 2) independent. The independent group consisted of variables like elevation, depth of water table,
potentiometric surface, glacial thickness, well depth, aquifer thickness, aquifer type, casing length, discharge, time, drawdown, well function, storativity; and dependent group consisted of transmissivity and hydraulic conductivity.

**Purpose and Objective**

The main purposes of this class project were to statistically analyze the dependent and independent variables of pumping well test data using principal component analysis (PCA) and cluster analysis and later depict the clusters spatially. The objective of using PCA was to reduce the number of considered variables and to replace large number of variables by a smaller number of variables while losing as little information as possible. Identification of the maximum variability in the data set linked with various combinations of variables. The objective of cluster analysis was to sort data into groups, or clusters, so that the degree of association was strong between members of the same cluster and weak between members of different clusters.

**Study Area**

Ground-water bearing formations within Geauga County (Figure 1) can be divided into two general groups: consolidated sandstone and shale formations within the bedrock, and unconsolidated glacial deposits. Of these two, the most wide-spread aquifers are within the Pennsylvanian sandstones of Pottsville Group, most prominently the Sharon Conglomerate and the Mississippian age Berea Sandstone (Table 1). The bedrock surface is deeply incised by buried valleys, forming a subglacial topography of ridges and hills consisting primarily of the more erosion resistive sandstones of Pennsylvanian system. Water wells’ yields from these formations may be in some locations as high as 545 m³/day (100 gpm) (Walker 1987). In the valleys and lowlands, the principal bedrock aquifers are the interbedded sandstone and shale of the Mississippian System (Figure 2 and 3). Unconsolidated aquifer formations are found primarily within the buried valleys. Outwash sand and gravel deposits in these valleys may yield more than 2725 m³/day (500 gallons per minute) from large diameter wells (Walker 1987). Other sand and gravel aquifers within the county include widely scattered kame deposits and alluvial deposits underlying the floodplains of some of the larger streams.

**Methodology**

All data were collected from Ohio Department of Natural Resources (ODNR) well data repository (Ohio Department of Natural Resources, n.d.). It contained around 80,000 wells within Geauga County alone. The data were collected randomly stratified meaning the data were collected without missing any individual street address within entire study area and collected 617 water well log and drilling reports.

Pumping tests are usually conducted only in public supply water wells, and there were very few of those in Geauga County, when compared with the number of single-home (residential) water wells. Each residential well was (supposed to be) tested for the production rate upon completion of drilling which consists of the following information:
1. Location
2. Construction details
3. Pumping test data (Static water level, rate of pumping, time, drawdown)
4. Well log

although the lithological descriptions are far from precise, they are adequate to distinguish between hydrostratigraphic units with varying permeability. Each lithological profile is defined macroscopically by a driller, i.e. not by a trained geologist. Nevertheless, drillers are well trained to distinguish macroscopically general types of sedimentary rock formations (Eckstein’s personal communication). All this data formed a basis for construction of the subsurface geology to reveal the aquifer thickness.

Aquifer properties are usually estimated by interpretation of constant-rate pumping test, during which the amount of drawdown is recorded at a given distance from the pumping well. Theis (1935) has shown the following relationship between the temporal and spatial distribution of drawdown around a pumping well as a function of the pumping rate and the hydraulic properties of an aquifer:

\[ h_0 - h(r,t) = \frac{Q}{4\pi T} \int_0^\infty \frac{e^{-y^2}}{y} dy = \frac{Q}{4\pi T} W(u) \]  

Where;
\( h_0 - h(r,t) \) is the drawdown [L]
\( t \) is the time since the pumping began [T]
\( r \) is the radial distance from the center of the pumped well to the observation well [L]
\( Q \) is the rate of pumping [L^3/T]
\( T \) is the coefficient of transmissibility [L^2/T]
\( K \) is the hydraulic conductivity [L/T]
\( b \) is the aquifer thickness [L]
\( u = r^2 S / 4Tt \)  \hspace{1cm} (2)

Where, \( S \) is storage coefficient of an aquifer (dimensionless)
\[ S = 4Ttu / r^2 \]  \hspace{1cm} (3)
\( (u) = -0.5772 - \ln u + u - u^2/2.2! + u^3/3.3! - u^4/4.4! + \cdots \)  \hspace{1cm} (4)

The drawdown observation was made within the pumped well \( r=r_w \). The \( r_w \) in residential water wells was very small (usually 3\(^\prime\) = 0.0762 m) \( r \) was even smaller, rendering the value of \( u \) (in eq. 2) small. Hence, only the first two terms on the right side of the eq. (4) remain significant. Following Cooper and Jacob (1946) and Jacob’s (1950) approximation, the eq. (1) can be simplified to

\[ h_0 - h(r_w,t) = \frac{Q}{4\pi T} (-0.5772 - \ln(u)) \]  \hspace{1cm} (5)

\[ T = k.b \]  \hspace{1cm} (6)

From the equations above, it can clearly be seen that hydraulic conductivity is dependent variable. It is directly link to transmissivity and thickness of aquifer and indirectly associated with pumping rate, storativity, drawdown, aquifer type, time, well radius and well function etc. The other variables like elevation, depth of water table, potentiometric surface, glacial deposit thickness, total well
depth, aquifer thickness, aquifer type, casing length, discharge, time, drawdown, well function, storativity, and township (location) are independent variables within the data set.

Once all the data had been collected, the data were plotted in different forms to understand nature of distribution of the data and to check and eliminate outliers, if any, from the data set.

**Principal Component Analysis**

PCA is a linear orthogonal transformation of data set that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate called the first principal component. The principal components are linear combinations of the original data (Everitt & Hothorn). It is a multivariate technique used to filter inter-correlation between variables. Its objectives were to transform the original variables into an identical number of uncorrelated or orthogonal components, i.e., Eigen vectors, and to short list Eigen vectors (2-4) that were independent (orthogonal) of one another but contain most of the covariance information in the original data (Haan, 1977). PCA generates an equal number of eigen vectors to the number of original variables. These eigen vectors have eigen values which measure their relative strength. An eigen vector with a high eigen value accounts for a large portion of the variance in the data. Each variable used in the analysis has a loading on each eigen vector. These loadings can be used to show the effect of the variables to the component. The sign of the loading indicates the relationship between each variable and the vector. Observations are also given scores which are analogous to the loadings. Observations with high scores on a variable are strongly influenced to the eigen vector.

PCA was selected for this study so as to eliminate inter-correlated variables losing as little information as possible and later identify the relationship between dependent and independent variables of pumping well test data. Altogether, there were eighteen variables in the data set. The variables were: elevation, depth of water table, potentiometric surface, glacial thickness, well depth, aquifer thickness, aquifer type, casing length, discharge, time, drawdown, well function, storativity, township, transmissivity and hydraulic conductivity. The variables were grouped into independent and dependent variables. It was assumed that dependent variables (transmissivity and hydraulic conductivity) were influenced by other variables in some degree.

PCA was run for multiple times to sort out the redundant variables. PCA was started with all the variables though some of them were known redundant because PCA is able to identify redundant variables. After running PCA for three times, the final data set was left with twelve variables.

**Cluster Analysis**

Cluster analysis is an exploratory data analysis tool for solving classification problems. Its objective is to sort data into groups, or clusters, so that the degree of association is strong between members of the same cluster and weak between members of different clusters. Each cluster thus describes, in terms of the data collected, the class to which its members belong (Romesburg, 2004).
Clustering was performed for the output data from final PCA run. Once clustering being done, it was now possible to plot the clusters in a space with help of different symbols. It will improve in the decision making on selection of productive groundwater zone and thus to install a public well or residential well.

Results

The distribution of Pumping test data, consisted of twenty variables, was shown in figure 4 but other variables like township, latitude, and longitude and aquifer type were not included in the figure. It was clearly seen that the most of the variables were normal in shape but five variables were not. So the non normal variables were later log transformed and their histograms were depicted and shown in figure 5. Later, Q-Q plot was portrayed in figure 6 to verify variables were normally distributed except for radius and storativity. The latter two had fixed values because all residential wells drilled with same radii and two different storativity values were chosen for two different types of aquifer in the study area. The steps in the diagram for Discharge and time of pumping were seen due to equal values used for large numbers of the samples in the data set. Box and whisker plot of variables from the dataset were depicted in figure 7. This diagram exposed outliers for some of the variables like casing length, elevation, glacial thickness, transmissivity etc. The nice thing about the diagram was its ability to clearly represent the distribution of dataset. Time of pumping, drawdown, casing length clearly demonstrated right skewed distribution whereas hydraulic conductivity, storativity, aquifer thickness has left or negative skewed distribution. The perfect bell curves were observed for well function, and transmissivity. The spatial distribution of hydraulic conductivity was plotted using ggplot 2 and illustrated in figure 8. Although, the plot could not describe clear spatial distribution, it was informative to see a general trend of data distribution over the study area. The plot would be much informative if different symbols for different weight of variables could be rendered in the diagram.

The correlation matrix created (Table 2) for first run of principal component analysis reflected perfect correlation between aquifer type and storativity value; and well function and ‘u’ component of well. Potentiometric surface and elevation revealed high correlation among others. On the other hand, water table depth had no correlation to township and radius. Moreover, the radius and time of pumping had very little effect on all other variables (Figure 14 and Table 2). So, the redundant variables needed to remove from the data set to correctly judge the effect of the one variable to another.

Studying the scree plot (Figure 9), it was concluded to choose five components as the most significant ones because those five components consisted of three fourth of the total variability in the data set (Figures 9). It can also be seen from eigen values of the components in the table 3.

The first component of PCA consisted of nearly one third of the total variability in the data (Table 3 & 4). The high loadings to the first component were coming from aquifer type, storativity, u, well function, and glacial thickness. Although this had higher variability than other, it needed to have closer look into it because most of the loadings were coming from redundant variables that were observed in the correlation matrix above. This component can be termed as well component since lot of variability coming from the well parameters. The PC2 had higher loadings contributing by water table
depth, well depth, aquifer thickness, drawdown, transmissivity, hydraulic conductivity. This component
is more related to the aquifer property and can be termed as aquifer component. Higher loading in PC3
were result of water table depth, potentiometric surface, glacial thickness, well depth, casing length,
transmissivity, well function and u component. It is a mixed signature from well and aquifer
characteristics. The PC4 was related to high values of loading associated with elevation, potentiometric
surface and well depth. This component can be named as z (elevation/depth) component. In PC5,
township, elevation, thickness, discharges, radius of well had high magnitude loadings. It was not clear
indication of any distinct component so can be termed as unknown component.

PCA Run 2

Two redundant variables, ‘u’ component of well and storativity, were eliminated from the first
set of data (Table 5). The variability of first component of the second run was decreased from one fourth
to one fifth of the total variability in the data set (Table 6 & 7). It was acceptable because elimination of
two high contributing variables was eliminated from the data set and aggregate variability was reduced.
The second PCA run illustrated elevation and potentiometric surface; and similarly transmissivity and
hydraulic conductivity had high correlation. Furthermore, water table depth had no correlation with
township and radius of the well. The time of pumping had low loadings in all components. Therefore,
time, elevation, township and radius of well were to eliminate from the data set for the next run.

The scree plot (Figure 10) obtained for second run also suggested to pick five components as
the significant ones because the summation of them provide only three fourth of the total variability in
the data set (Table 6 & 7).

The first component had high loadings characterized by casing length, aquifer type, and glacial
thickness. So it can be concluded that aquifer thickness had maximum variability in the data set. The
second component acquired its high loading from transmissivity, hydraulic conductivity, well depth and
drawdown which can be termed as aquifer component. The high loadings to the third component were
resulting from well depth, well function, drawdown and transmissivity. Hence this component can be
termed as well component. Elevation and potentiometric surface had maximum loadings for the fourth
component and can be termed as z component. For the fifth component, there was no distinct
indication of high loadings allocated from.

PCA Run 3

Once all the redundant and low scoring variables were eliminated from the data set (Table 8),
the third run had left with most prominent variables contributing to the variability so obtained. The third
run left with only twelve variables and only four components were chosen because it contributed almost
85% percentage of the total variability in the data set (Table 9 & 10).

The first component had higher loadings on water table depth, glacial deposit, well depth and
casing length. The variables were related to depth. Hence, it can be termed as depth component. Except
potentiometric surface and well function, all had negative loading to first component. The high loadings
to the second component coming from transmissivity, hydraulic conductivity, aquifer type, aquifer
thickness. Glacial thickness, aquifer type and casing length had negative loading and rest of the variables had positive loadings in the second component (Figure 13). The third component had higher loading on transmissivity, drawdown, aquifer type, well function and water table depth. The pumping rate was the only one most significant loading contributing to the fourth component. Transmissivity, hydraulic conductivity, water table depth and pumping rate had positive loading but other had negative loading. The score plot of two principal components were plotted to visually inspect the combination of the variables grouping in the data set (Figure 14). The higher the glacial thickness the lower the aquifer thickness and lower the hydraulic conductivity.

Cluster Analysis

After the PCA was done, cluster analysis undertaken to the final data set without any redundant and low scoring variables. The dendrogram plot was portrayed (Figure 12). For the analysis purpose, the dendrogram was grouped into four clusters at the height of 8. The four clusters consisted of 249, 88,143 and 34 sample points. Later the dendrogram was reshuffled to produce the number of samples contributing to each group from each township of the study area (Table 11). The table showed that maximum number of sample was contributed from Auburn Township and minimum from Thompson Township after eliminating the outliers from the dataset.

Yet another data was generated from the cluster analysis which provided the median value of each variable contributing to each group of cluster (Table 12). The data was very helpful to identify the combined effect of the variables to each group similar to the principal component analysis. This comprised four groups in it. Group 1 consisted data from high elevation with low glacial thickness. This group illustrated lower drawdown and higher hydraulic conductivity. The fourth group consisted of lower potentiometric surface, hydraulic conductivity and aquifer thickness but higher drawdown, and glacial thickness. The other two were not quite distinctive and needed further analysis. The total number of wells belonging to group 1 was much higher than the total number of wells pertaining to group 4.

Interpretation

Interpretation of multivariate dataset is very difficult task. The non normality of few variables in this data set was revealed due to the larger number of the observations in the narrow zone but the variable comprised the wide range of the data. The data had to be converted to normal distribution because we know behavior of the normal data distribution in much detail than any other distribution. Another set of non normal or semi-log normal data were observed due to the long range of data with discrete data set.

The interpretation of ggplot was difficult because of overlapping of the data points took place with same symbols and color. However, it gave a way to present spatial variation of data in the field. Upon improvement of this plot, valuable information could be achieved and helped in further understanding of the dataset.
The first component of principal component analysis under first run consisted of very high variability because of repeated loading attributed by similar variables. It was assumed that there were no any redundant variables in the third run of principal component analysis. The maximum loadings were contributed by water table depth, glacial depth, well depth and casing length. Hence it can be interpreted that thickness of aquifer had the greatest impact on well performance. As the glacial thickness increased, aquifer thickness slight increased and casing length increased proportionally but transmissivity and hydraulic conductivity decreased (Table 10, T and K are in negative log form initially) because the deeper aquifer (Berea Sandstone) had lower hydraulic conductivity. The larger variability was credited to second component by aquifer thickness, aquifer type, glacial thickness, drawdown, transmissivity and hydraulic conductivity. As the depth to the aquifer or water table increased, aquifer thickness decreased due to the glacial erosion leaving thin aquifer. Such aquifer had less hydraulic conductivity may be due to low water percolation into the aquifer through thick glacial till and hence produce higher drawdown.

The interpretation of cluster analysis leads to similar conclusion as of PCA. High elevation and low glacial thickness had lower drawdown and higher hydraulic conductivity. This might be due to the interconnection of the fracture to the wells and abundant of recharge through glacial till. The lower the potentiometric surface, hydraulic conductivity and aquifer thickness the higher the drawdown, and glacial thickness. This was due to presence of the buried valley covered with thick glacial deposits overlying the aquifer not allowing water to percolate into aquifer and lagging the recharge. The larger numbers of wells in group 1 indicated that the wells were productive and less number of well pertaining to group 4 illustrated the wells were not productive.

Summary

The pumping well data were analyzed using PCA and cluster analysis. Both the analyses looked for similar characteristics of the data. However, PCA was more robust than the cluster analysis though there are many varieties of cluster analyses to provide specific outputs. It is always helpful to perform PCA followed by cluster analysis. PCA deals with variables and can filter out redundant variables and remove false impression of high loading on the principal components. Whereas the cluster analysis makes cluster of individual sample points and helped to identify the clusters with individual data in it.

It can be concluded from all above analysis that the thickness of an aquifer had the most significant impact on the aquifer properties especially because the bed rock was fractured and underlain by glacial outwash. Water table depth (unconfined aquifer) and potentiometric surface (confined aquifer) lower as glacial thickness and well depth increased so as to produce lower hydraulic conductivity and higher drawdown. Group 4 was the least productive and group 1 was the most productive wells. This fact was also revealed from the numbers of wells belonging to the groups. Most of the wells belonging to group 1 were located in Auburn Township which emphasized Auburn Township was hydrogeologically the most productive within the study area.
References


Fuller, O.J., (1965) Bedrock geology of the Garrettsville Quadrangle; Ohio Department of Natural Resources, Division of Geological Survey, Report of Investigations No. 54, 26pp., 1 map.


Pedry, J. J., (1951) The geology of Chagrin Falls Township, Cuyahoga County and Bainbridge Township, Geauga County, Ohio; Ohio State University Master’s Thesis, Unpublished, 109 pp.


Theis, C. V. (1935). The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground-water storage. Transactions American Geophysical Union, 16, 519-524pp.

Walker, A.C. (1987) Ground-Water Resources of Geauga County; Ohio Department of Natural Resources, Division of Water, 1 map.
Table 1. Generalized bedrock stratigraphy of Geauga County, Ohio (modified from: Bower 1951; Fuller 1965; Pedry 1951; and Szmuc 1957).

<table>
<thead>
<tr>
<th>SYSTEM</th>
<th>GROUP</th>
<th>FORMATION</th>
<th>ROCK TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>PENNSYLVANIAN</td>
<td>Pottsville</td>
<td>Homewood Sandstone</td>
<td>White to tan, medium- to coarse-grained sandstone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mercer Shale</td>
<td>Blue-gray to black, silty to sandy, micaceous shale, locally interbedded with sandstone or siltstone layers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Massillon Sandstone</td>
<td>White to buff, fine- to medium-grained sandstone, locally contains thin layers of shale</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quakerstown Coal</td>
<td>Coal, found locally</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sharon Shale</td>
<td>Blue-gray, sandy micaceous shale, may contain thin layers of siltstone and/or sandstone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sharom Coal</td>
<td>Coal, found locally</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sharon Conglomerate</td>
<td>White, medium- to coarse-grained sandstone; contains lenses of pebbles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Meadville Shale</td>
<td>Interbedded blue-gray shales and siltstones</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sharpsville Sandstone</td>
<td>Blue-gray siltstone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Orangeville Shale</td>
<td>Blue-gray, silty shale with some siltstone layers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Berea Sandstone</td>
<td>Gray to blue-gray, fine- to medium-grained sandstone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bedford Shale</td>
<td>Blue-gray, silty shale with thin interbedded siltstones</td>
</tr>
<tr>
<td>DEVONIAN</td>
<td>Cuyahoga</td>
<td>Cleveland Shale</td>
<td>Dark-gray to black shale</td>
</tr>
<tr>
<td></td>
<td>Ohio Shale</td>
<td>Chagrin Shale</td>
<td>Blue-gray silty shale with some thin siltstone layers</td>
</tr>
</tbody>
</table>
Table 2. Correlation matrix of PC Analysis: Run 1

<table>
<thead>
<tr>
<th>Township Elevation</th>
<th>1</th>
<th>Township Depth</th>
<th>0</th>
<th>Potent. surface Depth</th>
<th>0.05</th>
<th>Glacial. th Depth</th>
<th>-0.11</th>
<th>Well. b</th>
<th>0.03</th>
<th>Aqu. type length</th>
<th>0.11</th>
<th>neg. Log.Q.</th>
<th>0.06</th>
<th>Log t</th>
<th>0.04</th>
<th>s</th>
<th>0.05</th>
<th>r</th>
<th>-0.05</th>
<th>neg. Log.T.</th>
<th>0.06</th>
<th>neg. Log.k.</th>
<th>0.12</th>
<th>W.u.</th>
<th>0.11</th>
<th>neg. Log.u.</th>
<th>0.11</th>
<th>S</th>
<th>-0.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Township Elevation</td>
<td>1</td>
<td>Township Depth</td>
<td>0</td>
<td>Potent. surface Depth</td>
<td>0.05</td>
<td>Glacial. th Depth</td>
<td>-0.11</td>
<td>Well. b</td>
<td>0.03</td>
<td>Aqu. type length</td>
<td>0.11</td>
<td>neg. Log.Q.</td>
<td>0.06</td>
<td>Log t</td>
<td>0.04</td>
<td>s</td>
<td>0.05</td>
<td>r</td>
<td>-0.05</td>
<td>neg. Log.T.</td>
<td>0.06</td>
<td>neg. Log.k.</td>
<td>0.12</td>
<td>W.u.</td>
<td>0.11</td>
<td>neg. Log.u.</td>
<td>0.11</td>
<td>S</td>
<td>-0.11</td>
</tr>
<tr>
<td>Township Elevation</td>
<td>1</td>
<td>Township Depth</td>
<td>0</td>
<td>Potent. surface Depth</td>
<td>0.05</td>
<td>Glacial. th Depth</td>
<td>-0.11</td>
<td>Well. b</td>
<td>0.03</td>
<td>Aqu. type length</td>
<td>0.11</td>
<td>neg. Log.Q.</td>
<td>0.06</td>
<td>Log t</td>
<td>0.04</td>
<td>s</td>
<td>0.05</td>
<td>r</td>
<td>-0.05</td>
<td>neg. Log.T.</td>
<td>0.06</td>
<td>neg. Log.k.</td>
<td>0.12</td>
<td>W.u.</td>
<td>0.11</td>
<td>neg. Log.u.</td>
<td>0.11</td>
<td>S</td>
<td>-0.11</td>
</tr>
</tbody>
</table>

The table above shows the correlation matrix for the PC Analysis of Run 1. The values in the table represent the correlation coefficients between different variables, with values ranging from -1 to 1. A value of 1 indicates a perfect positive correlation, while -1 indicates a perfect negative correlation. Values close to 0 indicate no correlation.
Table 3. Eigen values and Eigen vectors of component analysis run 1

<p>| | | | | | | | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>values</td>
<td>4.83</td>
<td>3.54</td>
<td>2.45</td>
<td>1.67</td>
<td>1.31</td>
<td>0.99</td>
<td>0.92</td>
<td>0.66</td>
<td>0.27</td>
<td>0.16</td>
<td>0.12</td>
<td>0.05</td>
<td>0.02</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.27</td>
<td>0.2</td>
<td>0.14</td>
<td>0.09</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.27</td>
<td>0.46</td>
<td>0.6</td>
<td>0.69</td>
<td>0.77</td>
<td>0.82</td>
<td>0.88</td>
<td>0.93</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vectors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Township</td>
<td>-0.08</td>
<td>-0.03</td>
<td>0.04</td>
<td>0.01</td>
<td>0.32</td>
<td>0.37</td>
<td>0.25</td>
<td>0.81</td>
<td>0.15</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Elevation</td>
<td>-0.2</td>
<td>0.01</td>
<td>0.17</td>
<td>0.63</td>
<td>0.23</td>
<td>0.02</td>
<td>0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.12</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.10</td>
<td>0.07</td>
<td>0.03</td>
<td>0.08</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>WT.Depth</td>
<td>-0.01</td>
<td>-0.3</td>
<td>0.04</td>
<td>0.14</td>
<td>0.14</td>
<td>0.01</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.61</td>
<td>0.43</td>
<td>0.21</td>
<td>0.11</td>
<td>0.04</td>
<td>0.04</td>
<td>0.20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Potent.surface</td>
<td>-0.19</td>
<td>0.1</td>
<td>0.29</td>
<td>0.56</td>
<td>0.18</td>
<td>0.02</td>
<td>0.03</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.11</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Glacial.th</td>
<td>0.33</td>
<td>0.14</td>
<td>0.17</td>
<td>0.05</td>
<td>0.08</td>
<td>0.08</td>
<td>0.04</td>
<td>0.06</td>
<td>0.13</td>
<td>0.62</td>
<td>0.31</td>
<td>0.46</td>
<td>0.15</td>
<td>0.06</td>
<td>0.03</td>
<td>0.06</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Well.Depth</td>
<td>0.1</td>
<td>0.38</td>
<td>0.3</td>
<td>0.25</td>
<td>0.13</td>
<td>0.13</td>
<td>0.01</td>
<td>0.01</td>
<td>0.24</td>
<td>0.16</td>
<td>0.24</td>
<td>0.07</td>
<td>0.7</td>
<td>0.10</td>
<td>0.06</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>-0.26</td>
<td>-0.27</td>
<td>0.04</td>
<td>0.09</td>
<td>0.25</td>
<td>0.21</td>
<td>0.10</td>
<td>0.05</td>
<td>0.6</td>
<td>0.28</td>
<td>0.02</td>
<td>0.09</td>
<td>0.48</td>
<td>0.21</td>
<td>0.03</td>
<td>0.08</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Aqu.type</td>
<td>0.4</td>
<td>0.14</td>
<td>0.14</td>
<td>0.12</td>
<td>0.07</td>
<td>0.05</td>
<td>0.00</td>
<td>0.05</td>
<td>0.24</td>
<td>0.02</td>
<td>0.09</td>
<td>0.10</td>
<td>0.03</td>
<td>0.12</td>
<td>0.42</td>
<td>0.07</td>
<td>0.69</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>casing.length</td>
<td>0.26</td>
<td>0.19</td>
<td>0.21</td>
<td>0.03</td>
<td>0.06</td>
<td>0.10</td>
<td>0.03</td>
<td>0.41</td>
<td>0.62</td>
<td>0.41</td>
<td>0.07</td>
<td>0.17</td>
<td>0.01</td>
<td>0.00</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>neg.Log.Q.</td>
<td>0.01</td>
<td>0.14</td>
<td>0.00</td>
<td>0.22</td>
<td>0.67</td>
<td>0.40</td>
<td>0.15</td>
<td>0.07</td>
<td>0.25</td>
<td>0.24</td>
<td>0.19</td>
<td>0.36</td>
<td>0.04</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>-0.06</td>
<td>0.16</td>
<td>0.11</td>
<td>0.13</td>
<td>0.18</td>
<td>0.65</td>
<td>0.54</td>
<td>0.14</td>
<td>0.3</td>
<td>0.02</td>
<td>0.11</td>
<td>0.12</td>
<td>0.04</td>
<td>0.07</td>
<td>0.23</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>0.06</td>
<td>0.36</td>
<td>0.27</td>
<td>0.03</td>
<td>0.33</td>
<td>0.06</td>
<td>0.24</td>
<td>0.16</td>
<td>0.25</td>
<td>0.14</td>
<td>0.26</td>
<td>0.66</td>
<td>0.03</td>
<td>0.09</td>
<td>0.07</td>
<td>0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.06</td>
<td>0.28</td>
<td>0.39</td>
<td>0.71</td>
<td>0.51</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>neg.Log.T.</td>
<td>-0.05</td>
<td>0.43</td>
<td>0.27</td>
<td>0.13</td>
<td>0.11</td>
<td>0.21</td>
<td>0.11</td>
<td>0.07</td>
<td>0.18</td>
<td>0.11</td>
<td>0.01</td>
<td>0.33</td>
<td>0.06</td>
<td>0.54</td>
<td>0.45</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>neg.Log.k.</td>
<td>-0.15</td>
<td>0.44</td>
<td>0.23</td>
<td>0.09</td>
<td>0.01</td>
<td>0.07</td>
<td>0.04</td>
<td>0.03</td>
<td>0.08</td>
<td>0.04</td>
<td>0.20</td>
<td>0.33</td>
<td>0.02</td>
<td>0.74</td>
<td>0.02</td>
<td>0.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>W.u.</td>
<td>-0.38</td>
<td>0.11</td>
<td>0.3</td>
<td>0.01</td>
<td>0.06</td>
<td>0.05</td>
<td>0.07</td>
<td>0.05</td>
<td>0.08</td>
<td>0.05</td>
<td>0.12</td>
<td>0.11</td>
<td>0.05</td>
<td>0.14</td>
<td>0.42</td>
<td>0.06</td>
<td>0.17</td>
<td>0.69</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>neg.Log.u.</td>
<td>-0.38</td>
<td>0.11</td>
<td>0.3</td>
<td>0.01</td>
<td>0.06</td>
<td>0.05</td>
<td>0.07</td>
<td>0.05</td>
<td>0.08</td>
<td>0.05</td>
<td>0.12</td>
<td>0.11</td>
<td>0.05</td>
<td>0.14</td>
<td>0.42</td>
<td>0.06</td>
<td>0.17</td>
<td>0.69</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.41</td>
<td>0.14</td>
<td>0.12</td>
<td>0.07</td>
<td>0.05</td>
<td>0.00</td>
<td>0.05</td>
<td>0.24</td>
<td>0.02</td>
<td>0.09</td>
<td>0.10</td>
<td>0.03</td>
<td>0.12</td>
<td>0.42</td>
<td>0.07</td>
<td>0.69</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Loading on PCA Run 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Township</td>
<td>-0.08</td>
<td>-0.03</td>
<td>0.04</td>
<td>0.01</td>
<td>0.32</td>
</tr>
<tr>
<td>Elevation</td>
<td>-0.2</td>
<td>0.01</td>
<td>0.17</td>
<td>0.63</td>
<td>-0.23</td>
</tr>
<tr>
<td>WT.Depth</td>
<td>0.01</td>
<td>-0.3</td>
<td>-0.43</td>
<td>0.14</td>
<td>-0.16</td>
</tr>
<tr>
<td>Potent.surface</td>
<td>-0.19</td>
<td>0.1</td>
<td>0.29</td>
<td>0.56</td>
<td>-0.18</td>
</tr>
<tr>
<td>Glacial.th</td>
<td>0.33</td>
<td>-0.14</td>
<td>-0.31</td>
<td>0.17</td>
<td>-0.05</td>
</tr>
<tr>
<td>Well.Depth</td>
<td>0.1</td>
<td>-0.38</td>
<td>-0.3</td>
<td>0.25</td>
<td>0.13</td>
</tr>
<tr>
<td>b</td>
<td>-0.26</td>
<td>-0.27</td>
<td>0.04</td>
<td>0.09</td>
<td>0.25</td>
</tr>
<tr>
<td>Aqu.type</td>
<td>0.41</td>
<td>0.14</td>
<td>0.14</td>
<td>0.12</td>
<td>0.07</td>
</tr>
<tr>
<td>casing.length</td>
<td>0.26</td>
<td>-0.19</td>
<td>-0.29</td>
<td>0.21</td>
<td>-0.03</td>
</tr>
<tr>
<td>neg.Log.Q.</td>
<td>0.01</td>
<td>-0.14</td>
<td>0</td>
<td>-0.22</td>
<td>-0.67</td>
</tr>
<tr>
<td>t</td>
<td>-0.06</td>
<td>0.16</td>
<td>-0.11</td>
<td>0.13</td>
<td>0.18</td>
</tr>
<tr>
<td>s</td>
<td>0.06</td>
<td>-0.36</td>
<td>0.27</td>
<td>0.03</td>
<td>0.33</td>
</tr>
<tr>
<td>r</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.03</td>
<td>-0.06</td>
<td>-0.28</td>
</tr>
<tr>
<td>neg.Log.T.</td>
<td>-0.05</td>
<td>-0.43</td>
<td>0.27</td>
<td>-0.13</td>
<td>-0.11</td>
</tr>
<tr>
<td>neg.Log.k.</td>
<td>-0.15</td>
<td>-0.44</td>
<td>0.23</td>
<td>-0.09</td>
<td>0.01</td>
</tr>
<tr>
<td>W.u.</td>
<td>-0.38</td>
<td>0.11</td>
<td>-0.3</td>
<td>-0.01</td>
<td>0.06</td>
</tr>
<tr>
<td>neg.Log.u.</td>
<td>-0.38</td>
<td>0.11</td>
<td>-0.3</td>
<td>-0.01</td>
<td>0.06</td>
</tr>
<tr>
<td>S</td>
<td>0.41</td>
<td>0.14</td>
<td>0.14</td>
<td>0.12</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Table 5. Correlation matrix of PCA Run 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Township</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elevation</td>
<td>0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT. Depth</td>
<td>0</td>
<td>0.03</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potent. surface</td>
<td>0.05</td>
<td>0.95</td>
<td>-0.27</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glacial. th</td>
<td>-0.11</td>
<td>-0.26</td>
<td>0.51</td>
<td>-0.4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well. Depth</td>
<td>-0.02</td>
<td>-0.02</td>
<td>0.72</td>
<td>-0.24</td>
<td>0.62</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>0.17</td>
<td>0.24</td>
<td>0.21</td>
<td>0.17</td>
<td>-0.35</td>
<td>0.41</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aqu. type</td>
<td>-0.11</td>
<td>-0.23</td>
<td>-0.24</td>
<td>-0.15</td>
<td>0.5</td>
<td>-0.01</td>
<td>-0.5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>casing. length</td>
<td>-0.1</td>
<td>-0.15</td>
<td>0.47</td>
<td>-0.28</td>
<td>0.76</td>
<td>0.61</td>
<td>-0.27</td>
<td>0.3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>neg. Log.Q.</td>
<td>-0.06</td>
<td>-0.04</td>
<td>0.23</td>
<td>-0.1</td>
<td>0.09</td>
<td>-0.01</td>
<td>-0.08</td>
<td>-0.09</td>
<td>0.03</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>0.04</td>
<td>0.07</td>
<td>-0.09</td>
<td>0.09</td>
<td>-0.03</td>
<td>-0.12</td>
<td>-0.09</td>
<td>0.04</td>
<td>-0.05</td>
<td>-0.08</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>0.05</td>
<td>-0.03</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.08</td>
<td>0.33</td>
<td>0.26</td>
<td>0.03</td>
<td>-0.15</td>
<td>-0.08</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>-0.05</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.02</td>
<td>0.05</td>
<td>-0.02</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>neg. Log.T.</td>
<td>0.06</td>
<td>0.04</td>
<td>0.16</td>
<td>-0.01</td>
<td>-0.06</td>
<td>0.22</td>
<td>0.3</td>
<td>-0.28</td>
<td>0.06</td>
<td>0.37</td>
<td>0.72</td>
<td>0.05</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>neg. Log.k.</td>
<td>0.12</td>
<td>0.13</td>
<td>0.2</td>
<td>0.06</td>
<td>-0.24</td>
<td>0.31</td>
<td>0.63</td>
<td>-0.45</td>
<td>-0.08</td>
<td>0.25</td>
<td>-0.19</td>
<td>0.65</td>
<td>0.04</td>
<td>0.9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>W.u.</td>
<td>0.11</td>
<td>0.21</td>
<td>0.14</td>
<td>0.16</td>
<td>-0.44</td>
<td>-0.12</td>
<td>0.31</td>
<td>-0.81</td>
<td>-0.32</td>
<td>-0.12</td>
<td>0.32</td>
<td>-0.38</td>
<td>-0.02</td>
<td>-0.26</td>
<td>-0.06</td>
<td>1</td>
</tr>
<tr>
<td>$values$</td>
<td>3.6</td>
<td>3.4</td>
<td>1.99</td>
<td>1.62</td>
<td>1.29</td>
<td>1</td>
<td>0.98</td>
<td>0.91</td>
<td>0.59</td>
<td>0.27</td>
<td>0.16</td>
<td>0.12</td>
<td>0.05</td>
<td>0.02</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>---</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Proportion</td>
<td>0.23</td>
<td>0.21</td>
<td>0.12</td>
<td>0.1</td>
<td>0.08</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cumulating</td>
<td>0.23</td>
<td>0.44</td>
<td>0.56</td>
<td>0.66</td>
<td>0.74</td>
<td>0.81</td>
<td>0.87</td>
<td>0.92</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Township</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.3</td>
<td>-0.4</td>
<td>0.05</td>
<td>0.84</td>
<td>-0.1</td>
<td>0.02</td>
<td>-0</td>
<td>0.01</td>
<td>0.04</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
</tr>
<tr>
<td>Elevation</td>
<td>-0.3</td>
<td>-0.1</td>
<td>0.08</td>
<td>-0.6</td>
<td>0.21</td>
<td>-0</td>
<td>0.05</td>
<td>0.04</td>
<td>-0</td>
<td>0.1</td>
<td>0.04</td>
<td>-0</td>
<td>0.05</td>
<td>-0.1</td>
<td>0.04</td>
<td>-0.2</td>
</tr>
<tr>
<td>WT.Depth</td>
<td>0.23</td>
<td>-0.3</td>
<td>0.47</td>
<td>-0</td>
<td>0.16</td>
<td>-0</td>
<td>0.05</td>
<td>0.05</td>
<td>0</td>
<td>-0.6</td>
<td>0.43</td>
<td>0.2</td>
<td>-0.1</td>
<td>0.03</td>
<td>0.07</td>
<td>0.19</td>
</tr>
<tr>
<td>Potent.surface</td>
<td>-0.3</td>
<td>-0</td>
<td>-0.1</td>
<td>-0.6</td>
<td>0.16</td>
<td>-0</td>
<td>0.03</td>
<td>0.03</td>
<td>-0</td>
<td>0.07</td>
<td>-0.1</td>
<td>0</td>
<td>-0.1</td>
<td>0.04</td>
<td>0.21</td>
<td>0.67</td>
</tr>
<tr>
<td>Glacial.th</td>
<td>0.47</td>
<td>0.03</td>
<td>0.16</td>
<td>-0.2</td>
<td>0.04</td>
<td>-0.1</td>
<td>-0</td>
<td>0.06</td>
<td>0</td>
<td>-0.2</td>
<td>-0.6</td>
<td>-0.3</td>
<td>-0.5</td>
<td>-0.2</td>
<td>0.03</td>
<td>-0</td>
</tr>
<tr>
<td>Well.Depth</td>
<td>0.3</td>
<td>-0.3</td>
<td>0.31</td>
<td>-0.2</td>
<td>-0.1</td>
<td>0.12</td>
<td>0.03</td>
<td>-0</td>
<td>0.25</td>
<td>0.14</td>
<td>-0.2</td>
<td>-0.1</td>
<td>0.7</td>
<td>0.11</td>
<td>-0</td>
<td>0.07</td>
</tr>
<tr>
<td>$b$</td>
<td>-0.2</td>
<td>-0.4</td>
<td>0.15</td>
<td>0.01</td>
<td>-0.2</td>
<td>0.19</td>
<td>0.12</td>
<td>0</td>
<td>0.57</td>
<td>0.27</td>
<td>0.03</td>
<td>0.1</td>
<td>-0.5</td>
<td>0.2</td>
<td>0.02</td>
<td>-0.1</td>
</tr>
<tr>
<td>Aqu.type</td>
<td>0.31</td>
<td>0.27</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0</td>
<td>0.08</td>
<td>0.41</td>
<td>-0.1</td>
<td>0.21</td>
<td>0.1</td>
<td>-0</td>
<td>-0.1</td>
<td>-0.6</td>
<td>0.2</td>
</tr>
<tr>
<td>casing.length</td>
<td>0.41</td>
<td>-0.1</td>
<td>0.16</td>
<td>-0.2</td>
<td>0.01</td>
<td>-0.1</td>
<td>-0.1</td>
<td>0</td>
<td>-0.4</td>
<td>0.64</td>
<td>0.38</td>
<td>0.06</td>
<td>-0.2</td>
<td>0.01</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>neg.Log.Q.</td>
<td>0.06</td>
<td>-0.1</td>
<td>-0</td>
<td>0.21</td>
<td>0.7</td>
<td>-0.4</td>
<td>0.05</td>
<td>0.06</td>
<td>0.25</td>
<td>0.23</td>
<td>-0.2</td>
<td>0.36</td>
<td>0.04</td>
<td>-0.1</td>
<td>0.01</td>
<td>-0</td>
</tr>
<tr>
<td>$t$</td>
<td>-0.1</td>
<td>0.13</td>
<td>0.16</td>
<td>-0.1</td>
<td>-0.2</td>
<td>-0.5</td>
<td>-0.7</td>
<td>-0.2</td>
<td>0.27</td>
<td>-0</td>
<td>0.11</td>
<td>-0.1</td>
<td>0.03</td>
<td>0.04</td>
<td>0.17</td>
<td>-0</td>
</tr>
<tr>
<td>$s$</td>
<td>0.12</td>
<td>-0.4</td>
<td>-0.3</td>
<td>-0.1</td>
<td>-0.3</td>
<td>0</td>
<td>-0.2</td>
<td>-0.1</td>
<td>-0.2</td>
<td>-0.2</td>
<td>-0.2</td>
<td>0.66</td>
<td>-0</td>
<td>-0.1</td>
<td>0.04</td>
<td>-0</td>
</tr>
<tr>
<td>$r$</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>0.05</td>
<td>0.29</td>
<td>0.59</td>
<td>-0.6</td>
<td>0.44</td>
<td>0.04</td>
<td>0.01</td>
<td>-0</td>
<td>0.03</td>
<td>-0</td>
<td>0</td>
<td>-0</td>
<td>0</td>
</tr>
<tr>
<td>neg.Log.T.</td>
<td>0.04</td>
<td>-0.5</td>
<td>-0.3</td>
<td>0.07</td>
<td>0.14</td>
<td>-0.2</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.2</td>
<td>-0.1</td>
<td>0</td>
<td>-0.4</td>
<td>-0.1</td>
<td>0.6</td>
<td>-0.3</td>
<td>0.06</td>
</tr>
<tr>
<td>neg.Log.k.</td>
<td>-0.1</td>
<td>-0.5</td>
<td>-0.2</td>
<td>0.08</td>
<td>0.02</td>
<td>-0</td>
<td>-0.1</td>
<td>-0</td>
<td>0.05</td>
<td>0.04</td>
<td>0.19</td>
<td>-0.4</td>
<td>0.04</td>
<td>-0.7</td>
<td>-0</td>
<td>0.09</td>
</tr>
<tr>
<td>W.u.</td>
<td>-0.3</td>
<td>-0</td>
<td>0.49</td>
<td>0.17</td>
<td>-0.1</td>
<td>-0</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.3</td>
<td>0.09</td>
<td>-0.2</td>
<td>0.13</td>
<td>-0.1</td>
<td>-0.1</td>
<td>-0.6</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Table 7. Loading on PCA Run 2

<table>
<thead>
<tr>
<th>$values$</th>
<th>3.6</th>
<th>3.4</th>
<th>1.99</th>
<th>1.62</th>
<th>1.29</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion</td>
<td>0.23</td>
<td>0.21</td>
<td>0.12</td>
<td>0.1</td>
<td>0.08</td>
</tr>
<tr>
<td>cumulating</td>
<td>0.23</td>
<td>0.44</td>
<td>0.56</td>
<td>0.66</td>
<td>0.74</td>
</tr>
</tbody>
</table>

$vectors$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Township</td>
<td>-0.09</td>
<td>-0.08</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.3</td>
</tr>
<tr>
<td>Elevatio</td>
<td>-0.27</td>
<td>-0.1</td>
<td>0.08</td>
<td>-0.6</td>
<td>0.21</td>
</tr>
<tr>
<td>WT.Depth</td>
<td>0.23</td>
<td>-0.25</td>
<td>0.47</td>
<td>-0</td>
<td>0.16</td>
</tr>
<tr>
<td>Potent.surface</td>
<td>-0.3</td>
<td>-0.02</td>
<td>-0.06</td>
<td>-0.6</td>
<td>0.16</td>
</tr>
<tr>
<td>Glacial.th</td>
<td>0.47</td>
<td>0.03</td>
<td>0.16</td>
<td>-0.2</td>
<td>0.04</td>
</tr>
<tr>
<td>Well.Depth</td>
<td>0.3</td>
<td>-0.3</td>
<td>0.31</td>
<td>-0.2</td>
<td>-0.14</td>
</tr>
<tr>
<td>b</td>
<td>-0.19</td>
<td>-0.4</td>
<td>0.15</td>
<td>0.01</td>
<td>-0.24</td>
</tr>
<tr>
<td>Aqu.type</td>
<td>0.31</td>
<td>0.27</td>
<td>-0.3</td>
<td>-0.2</td>
<td>-0.08</td>
</tr>
<tr>
<td>casing.length</td>
<td>0.41</td>
<td>-0.05</td>
<td>0.16</td>
<td>-0.2</td>
<td>0.01</td>
</tr>
<tr>
<td>neg.Log.Q.</td>
<td>0.06</td>
<td>-0.12</td>
<td>-0.04</td>
<td>0.21</td>
<td>0.7</td>
</tr>
<tr>
<td>t</td>
<td>-0.09</td>
<td>0.13</td>
<td>0.16</td>
<td>-0.1</td>
<td>-0.17</td>
</tr>
<tr>
<td>s</td>
<td>0.12</td>
<td>-0.4</td>
<td>-0.3</td>
<td>-0.1</td>
<td>-0.3</td>
</tr>
<tr>
<td>r</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.04</td>
<td>0.05</td>
<td>0.29</td>
</tr>
<tr>
<td>neg.Log.T.</td>
<td>0.04</td>
<td>-0.5</td>
<td>-0.3</td>
<td>0.07</td>
<td>0.14</td>
</tr>
<tr>
<td>neg.Log.k.</td>
<td>-0.07</td>
<td>-0.5</td>
<td>-0.18</td>
<td>0.08</td>
<td>0.02</td>
</tr>
<tr>
<td>W.u.</td>
<td>-0.3</td>
<td>-0.01</td>
<td>0.49</td>
<td>0.17</td>
<td>-0.06</td>
</tr>
</tbody>
</table>

Table 10. Loading on PCA Run 3

<table>
<thead>
<tr>
<th>$values$</th>
<th>3.41</th>
<th>3.28</th>
<th>1.96</th>
<th>1.27</th>
</tr>
</thead>
<tbody>
<tr>
<td>proportion</td>
<td>0.28</td>
<td>0.27</td>
<td>0.16</td>
<td>0.11</td>
</tr>
<tr>
<td>cumulative</td>
<td>0.28</td>
<td>0.56</td>
<td>0.72</td>
<td>0.83</td>
</tr>
</tbody>
</table>

$vectors$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>WT.Depth</td>
<td>0.35</td>
<td>-0.07</td>
<td>-0.5</td>
<td>0.11</td>
</tr>
<tr>
<td>Potent.surface</td>
<td>-0.22</td>
<td>-0.14</td>
<td>0.11</td>
<td>-0.2</td>
</tr>
<tr>
<td>Glacial.th</td>
<td>0.41</td>
<td>0.31</td>
<td>-0.12</td>
<td>0</td>
</tr>
<tr>
<td>Well.Depth</td>
<td>0.44</td>
<td>-0.07</td>
<td>-0.27</td>
<td>-0.3</td>
</tr>
<tr>
<td>b</td>
<td>0.04</td>
<td>-0.4</td>
<td>-0.15</td>
<td>-0.3</td>
</tr>
<tr>
<td>Aqu.type</td>
<td>0.12</td>
<td>0.42</td>
<td>0.35</td>
<td>-0.1</td>
</tr>
<tr>
<td>casing.length</td>
<td>0.4</td>
<td>0.21</td>
<td>-0.12</td>
<td>-0.1</td>
</tr>
<tr>
<td>neg.Log.Q.</td>
<td>0.11</td>
<td>-0.07</td>
<td>0.02</td>
<td>0.81</td>
</tr>
<tr>
<td>s</td>
<td>0.29</td>
<td>-0.24</td>
<td>0.38</td>
<td>-0.3</td>
</tr>
<tr>
<td>neg.Log.T.</td>
<td>0.27</td>
<td>-0.4</td>
<td>0.31</td>
<td>0.22</td>
</tr>
<tr>
<td>neg.Log.k.</td>
<td>0.21</td>
<td>-0.5</td>
<td>0.18</td>
<td>0.07</td>
</tr>
<tr>
<td>W.u.</td>
<td>-0.27</td>
<td>-0.21</td>
<td>-0.5</td>
<td>-0</td>
</tr>
</tbody>
</table>
Table 8. Correlation matrix for PCA Run 3

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>WT. Depth</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potent. surface</td>
<td>-0.27</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glacial. th</td>
<td>0.51</td>
<td>-0.4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well. Depth</td>
<td>0.72</td>
<td>-0.24</td>
<td>0.62</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>0.21</td>
<td>0.17</td>
<td>-0.35</td>
<td>0.41</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aqu. type</td>
<td>-0.24</td>
<td>-0.15</td>
<td>0.5</td>
<td>-0.01</td>
<td>-0.5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>casing.length</td>
<td>0.47</td>
<td>-0.28</td>
<td>0.76</td>
<td>0.61</td>
<td>-0.27</td>
<td>0.3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>neg. Log.Q.</td>
<td>0.23</td>
<td>-0.1</td>
<td>0.09</td>
<td>-0.01</td>
<td>-0.08</td>
<td>-0.09</td>
<td>0.03</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>0.02</td>
<td>-0.03</td>
<td>0.08</td>
<td>0.33</td>
<td>0.26</td>
<td>0.03</td>
<td>0.2</td>
<td>-0.15</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>neg. Log.T.</td>
<td>0.16</td>
<td>-0.01</td>
<td>-0.06</td>
<td>0.22</td>
<td>0.3</td>
<td>-0.28</td>
<td>0.06</td>
<td>0.37</td>
<td>0.72</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>neg. Log.k.</td>
<td>0.2</td>
<td>0.06</td>
<td>-0.24</td>
<td>0.31</td>
<td>0.63</td>
<td>-0.45</td>
<td>-0.08</td>
<td>0.25</td>
<td>0.65</td>
<td>0.9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>W.u.</td>
<td>0.14</td>
<td>0.16</td>
<td>-0.44</td>
<td>-0.12</td>
<td>0.31</td>
<td>-0.81</td>
<td>-0.32</td>
<td>-0.12</td>
<td>-0.38</td>
<td>-0.26</td>
<td>-0.06</td>
<td>1</td>
</tr>
</tbody>
</table>
### Table 9. Eigen values and Eigen vectors PCA Run 3

<table>
<thead>
<tr>
<th></th>
<th>values</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3.41</td>
<td>3.28</td>
<td>1.96</td>
<td>1.27</td>
<td>0.8</td>
<td>0.62</td>
<td>0.26</td>
<td>0.17</td>
<td>0.12</td>
<td>0.05</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>proportion</td>
<td>0.28</td>
<td>0.27</td>
<td>0.16</td>
<td>0.11</td>
<td>0.07</td>
<td>0.05</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>cumulative</td>
<td>0.28</td>
<td>0.56</td>
<td>0.72</td>
<td>0.83</td>
<td>0.89</td>
<td>0.94</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>WT.Depth</td>
<td>0.35</td>
<td>-0.07</td>
<td>-0.46</td>
<td>0.11</td>
<td>-0.11</td>
<td>0.02</td>
<td>0.62</td>
<td>0.39</td>
<td>0.31</td>
<td>-0.1</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Potent.surface</td>
<td>-0.22</td>
<td>-0.14</td>
<td>0.11</td>
<td>-0.2</td>
<td>-0.92</td>
<td>-0.14</td>
<td>0.11</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glacial.th</td>
<td>0.41</td>
<td>0.31</td>
<td>-0.12</td>
<td>0</td>
<td>-0.1</td>
<td>-0.03</td>
<td>0.17</td>
<td>-0.52</td>
<td>-0.42</td>
<td>-0.44</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Well.Dpth</td>
<td>0.44</td>
<td>-0.07</td>
<td>-0.27</td>
<td>-0.25</td>
<td>-0.12</td>
<td>0.23</td>
<td>-0.14</td>
<td>-0.21</td>
<td>-0.13</td>
<td>0.69</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b</td>
<td>0.04</td>
<td>-0.43</td>
<td>-0.15</td>
<td>-0.28</td>
<td>-0.03</td>
<td>0.58</td>
<td>-0.28</td>
<td>-0.02</td>
<td>0.11</td>
<td>-0.45</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aqu.type</td>
<td>0.12</td>
<td>0.42</td>
<td>0.35</td>
<td>-0.11</td>
<td>-0.1</td>
<td>0.37</td>
<td>0.05</td>
<td>0.06</td>
<td>0.24</td>
<td>-0.18</td>
<td>-0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>casing.length</td>
<td>0.4</td>
<td>0.21</td>
<td>-0.12</td>
<td>-0.08</td>
<td>-0.16</td>
<td>-0.42</td>
<td>-0.62</td>
<td>0.36</td>
<td>0.14</td>
<td>-0.18</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>neg.Log.Q.</td>
<td>0.11</td>
<td>-0.07</td>
<td>0.02</td>
<td>0.81</td>
<td>-0.24</td>
<td>-0.2</td>
<td>-0.23</td>
<td>-0.27</td>
<td>0.32</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>s</td>
<td>0.29</td>
<td>-0.24</td>
<td>0.38</td>
<td>-0.29</td>
<td>0.14</td>
<td>-0.28</td>
<td>0.13</td>
<td>-0.38</td>
<td>0.58</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>neg.Log.T.</td>
<td>0.27</td>
<td>-0.37</td>
<td>0.31</td>
<td>0.22</td>
<td>0.02</td>
<td>-0.23</td>
<td>0.11</td>
<td>0.04</td>
<td>-0.31</td>
<td>-0.1</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>neg.Log.k.</td>
<td>0.21</td>
<td>-0.48</td>
<td>0.18</td>
<td>0.07</td>
<td>0.03</td>
<td>0.04</td>
<td>-0.04</td>
<td>0.22</td>
<td>-0.27</td>
<td>-0.07</td>
<td>-0.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W.u.</td>
<td>-0.27</td>
<td>-0.21</td>
<td>-0.51</td>
<td>-0.02</td>
<td>0.09</td>
<td>-0.31</td>
<td>-0.09</td>
<td>-0.35</td>
<td>0.14</td>
<td>-0.18</td>
<td>-0.56</td>
</tr>
</tbody>
</table>

### Table 11. Sample point distribution in each group

| groups | Auburn | Bainbridge | Burton | Chardon | Ches t er | Clari don | Hamb den | Hunts burg | Middle field | Mont ville | Mun son | Newbury | Park man | Russell | Thompson | Troy | Total |
|--------|--------|------------|--------|---------|-----------|-----------|----------|------------|--------------|------------|----------|---------|---------|----------|---------|----------|------|-------|
| 1      | 60     | 23         | 10     | 6       | 12        | 15        | 12       | 8          | 13           | 7          | 23       | 22      | 8       | 15       | 5       | 10     | 249  |
| 2      | 10     | 12         | 3      | 5       | 27        | 1         | 4        | 1          | 4            | 2          | 7        | 3       | 1       | 8        | 0       | 0      | 88   |
| 3      | 34     | 29         | 14     | 11      | 8         | 6         | 3        | 8          | 8            | 8          | 38       | 19      | 16      | 30       | 2       | 9      | 143  |
| 4      | 1      | 10         | 1      | 3       | 3         | 0         | 0        | 1          | 0            | 1          | 6        | 1       | 1       | 5        | 1       | 0      | 34   |
Table 12. Median values of each variable in corresponding groups

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.62</td>
<td>353.81</td>
<td>13.41</td>
<td>24.38</td>
<td>10.67</td>
<td>2.00</td>
<td>13.72</td>
<td>2.90</td>
<td>3.68</td>
<td>4.65</td>
<td>6.10</td>
<td>8.41</td>
</tr>
<tr>
<td>2</td>
<td>21.19</td>
<td>319.32</td>
<td>26.52</td>
<td>39.62</td>
<td>12.04</td>
<td>2.00</td>
<td>24.69</td>
<td>3.02</td>
<td>3.83</td>
<td>4.80</td>
<td>4.57</td>
<td>8.51</td>
</tr>
<tr>
<td>3</td>
<td>13.41</td>
<td>354.90</td>
<td>8.23</td>
<td>30.48</td>
<td>19.20</td>
<td>1.00</td>
<td>11.58</td>
<td>3.02</td>
<td>4.12</td>
<td>5.39</td>
<td>7.62</td>
<td>11.79</td>
</tr>
<tr>
<td>4</td>
<td>16.31</td>
<td>318.14</td>
<td>25.61</td>
<td>41.15</td>
<td>13.42</td>
<td>2.00</td>
<td>25.91</td>
<td>3.02</td>
<td>4.36</td>
<td>5.44</td>
<td>22.86</td>
<td>6.67</td>
</tr>
</tbody>
</table>
Figure 1. Location of study area

Figure 2. Conceptual model of the bedrock formations and a buried valley in Geauga County, Ohio (adapted from Seyoum and Eckstein, 2012)
Figure 3. Fracture styles in Sharon Conglomerate, Meadville Shale, Sharpsville Sandstone, Orangeville Shale (undifferentiated), and Berea Sandstone.
Figure 4. Histogram of initial data
Figure 5. Histogram of log transformation of initial non normal data

- Hydraulic conductivity
- Discharge
- Transmissivity
- Time of pumping
Figure 6. Q-Q plot of pump test variables
Figure 7. Box and whisker plot of pumping well test data
Figure 8. ggplot2 for spatial distribution of the hydraulic conductivity
Figure 9. Scree plot for PCA Run 1

Figure 10. Scree plot for PCA Run 2

Figure 11. Scree plot for PCA Run 3
Figure 12. Dendrogram of pumping test data
Figure 13. Loadings to the four most variable components
Figure 14. Score plot of principal components analysis run 1
R code

Histogram plot (histinitial.csv)
attach(histdata)
par(mfrow = c(4,4) )
hist(b, xlab="aquifer thickness", main="")
hist(casing.length, xlab="casing length", main="")
hist(Elevation..m., xlab="Elevation", main="")
hist(Glacial.th, xlab="Glacial thickness", main="")
hist(k, xlab="hydraulic conductivity", main="")
hist(Potent.surface, xlab="Potential surface", main="")
hist(Q, xlab="Discharge", main="")
hist(r, xlab="radius", main="")
hist(s, xlab="drawdown", main="")
hist(S, xlab="Storativity", main="")
hist(t, xlab="time of pumping", main="")
hist(T, xlab="Transmissivity", main="")
hist(u, xlab="u", main="")
hist(W.u., xlab="Well function", main="")
hist(Well.Depth, xlab="Well Depth", main="")
hist(WT.Depth, xlab="Water table", main="")

Histogram plot (Log transformed initial data.csv)
attach(histdata)
par(mfrow=c(3,2))
hist(histdata$neg.log.k, xlab="hydraulic conductivity", main="")
hist(histdata$neg.log.Q, xlab="Discharge", main="")
hist(histdata$neg.log.u, xlab="u", main="")
hist(histdata$neg.log.T, xlab="Transmissivity", main="")
hist(histdata$log.t, xlab="time of pumping", main="")

Q-Q plot (Log transformed initial data.csv)
attach(qqdata)
par(mfrow = c(4,4) )
qqnorm (qqdata$b, xlab="aquifer thickness", main="", ylab="normdist(x)"); qqline(qqdata$b, col=2)
qqnorm(qqdata$casing.length, xlab="casing length", main="", ylab="normdist(x)");
qqline(qqdata$casing.length, col=2)
qqnorm(qqdata$Elevation, xlab="Elevation", main="", ylab="normdist(x)"); qqline(qqdata$Elevation, col=2)
qqnorm(qqdata$Glacial.th, xlab="Glacial thickness", main="", ylab="normdist(x)"); qqline(qqdata$Glacial.th, col=2)
qqnorm(qqdata$neg.log.k, xlab="hydraulic conductivity", main="", ylab="normdist(x)");
qqline(qqdata$neg.log.k, col=2)
qqnorm(qqdata$Potent.surface, xlab="Potential surface", main="", ylab="normdist(x) ");
qqline(qqdata$ Potent.surface, col=2)
qqnorm(qqdata$neg.log.Q, xlab="Discharge", main="", ylab="normdist(x) ");
qqline(qqdata$ neg.log.Q, col=2)
qqnorm(qqdata$r, xlab="radius", main="", ylab="normdist(x) ");
qqline(qqdata$r , col=2)
qqnorm(qqdata$s, xlab="drawdown", main="", ylab="normdist(x) ");
qqline(qqdata$s , col=2)
qqnorm(qqdata$S, xlab="Storativity", main="", ylab="normdist(x) ");
qqline(qqdata$ S, col=2)
qqnorm(qqdata$log.t, xlab="time of pumping", main="", ylab="normdist(x) ");
qqline(qqdata$ log.t, col=2)
qqnorm(qqdata$neg.log.T, xlab="transmissivity", main="", ylab="normdist(x) ");
qqline(qqdata$neg.log.T , col=2)
qqnorm(qqdata$neg.log.u, xlab="neg.log.u", main="", ylab="normdist(x) ");
qqline(qqdata$neg.log.u, col=2)
qqnorm(qqdata$W.u, xlab="Well function", main="", ylab="normdist(x) ");
qqline(qqdata$W.u , col=2)
qqnorm(qqdata$Well.Depth, xlab="Well Depth", main="", ylab="normdist(x) ");
qqline(qqdata$Well.Depth, col=2)

Box plot (log transformed outlier eliminated.csv)
attach(boxdata)
par(mfrow = c(4,4) )
par(mar= c(4,3,1,1))
boxplot(boxdata$b, xlab="aquifer thickness", main="")
boxplot(boxdata$casing.length, xlab="casing length", main="")
boxplot(boxdata$Elevation.m., xlab="Elevation", main=""
boxplot(boxdata$Glacial.th, xlab="Glacial thickness", main=""
boxplot(boxdata$neg.log.k, xlab="hydraulic conductivity", main=""
boxplot(boxdata$Potent.surface, xlab="Potential surface", main=""
boxplot(boxdata$neg.log.Q, xlab="Discharge", main=""
boxplot(boxdata$r, xlab="radius", main=""
boxplot(boxdata$s, xlab="drawdown", main=""
boxplot(boxdata$S, xlab="Storativity", main=""
boxplot(boxdata$log.t, xlab="time of pumping", main=""
boxplot(boxdata$neg.log.T, xlab="transmissivity", main=""
boxplot(boxdata$neg.log.u, xlab="u", main=""
boxplot(boxdata$W.u, xlab="Well function", main=""
boxplot(boxdata$Well.Depth, xlab="Well Depth", main=""
boxplot(boxdata$WT.Depth, xlab="Water table", main=""

GGplot2 (Log transformed initial data.csv)
Library(ggplot2)
Load(ggplot2)
qplot(data = bubbledata, Longitude,Latitude,size = neg.log.k, main="ggplot of hydraulic conductivity and its spatial distribution")

**PCA (rdataPCA1.csv)**

```r
#PCA (rdataPCA1.csv)

rdataPCA1=read.csv("C:\\Users\\GEO.Madan\\Dropbox\\Rclass project\\rdataPCA1.csv", header=T)
R=round(cor(rdataPCA1),digits=2)
myEig =eigen(R)
rdataPCA1.pc= prcomp(rdataPCA1, retx=TRUE, center=TRUE,scale.=TRUE)
sd1<- myEig$values
loadings <- rdataPCA1.pc$rotation
rownames(loadings) <- colnames(rdataPCA1)
scores <- rdataPCA1.pc$x
# scree plot
plot(sd1,xlab="component number" ,ylab="eigenvalue",pch=12,cex=1.3)

par(mfrow=c(2,2)); par(mar=c(4,4,1,1))
# this plots the eigenvectors
plot(scores[,1], scores[,2], xlab="PCA 1", ylab="PCA 2",type="n",xlim=c(-6,5),ylim=c(-6,3))
arrows(0,0,loadings[,1]*10,loadings[,2]*10, length=0.1,angle=20, col="red")
#1.2 is the measure of the position of text away from the arrow point
text(loadings[,1]*10*1.2,loadings[,2]*10*1.2,rownames(loadings), col="blue", cex=0.7)

# this plots the eigenvectors
plot(scores[,1], scores[,3], xlab="PCA 1", ylab="PCA 3",type="n",xlim=c(-6,5),ylim=c(-6,3))
arrows(0,0,loadings[,1]*10,loadings[,3]*10, length=0.1,angle=20, col="red")
#1.2 is the measure of the position of text away from the arrow point
text(loadings[,1]*10*1.2,loadings[,3]*10*1.2,rownames(loadings), col="blue", cex=0.7)

# this plots the eigenvectors
plot(scores[,2], scores[,3], xlab="PCA 2", ylab="PCA 3",type="n",xlim=c(-6,3),ylim=c(-4,5))
arrows(0,0,loadings[,2]*10,loadings[,3]*10, length=0.1,angle=20, col="red")
#1.2 is the measure of the position of text away from the arrow point
text(loadings[,2]*10*1.2,loadings[,3]*10*1.2,rownames(loadings), col="blue", cex=0.7)

# this plots the eigenvectors
plot(scores[,1], scores[,4], xlab="PCA 1", ylab="PCA 4",type="n",xlim=c(-6,4),ylim=c(-6,3))
arrows(0,0,loadings[,1]*10,loadings[,4]*10, length=0.1,angle=20, col="red")
#1.2 is the measure of the position of text away from the arrow point
text(loadings[,1]*10*1.2,loadings[,4]*10*1.2,rownames(loadings), col="blue", cex=0.7)
```
```r
loadings1 <- read.csv("C:\\Users\\GEO.Madan\\Dropbox\\Rclass project\\loadings1.csv", header=T)
par(mfrow= c(2,2))
barplot(loadings1$PC1, main="barplot of component 1", xlab="component 1", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
barplot(loadings1$PC2, main="barplot of component 2", xlab="component 2", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
barplot(loadings1$PC3, main="barplot of component 3", xlab="component 3", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
barplot(loadings1$PC4, main="barplot of component 4", xlab="component 4", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)

# biplot is a built in function to plot both scores and loadings at the same time simultaneously
biplot(scores[,1:2], loadings[,1:2], xlab=rownames(scores),ylab=NULL, cex=0.4,xlim=c(-3,3),ylim=c(-3,3),var.axes=TRUE)

--------------------------------------------------------------------------------------------------------------------------
# PCA (rdataPCA2.csv)
rdataPCA2 = read.csv("C:\\Users\\GEO.Madan\\Dropbox\\Rclass project\\rdataPCA2.csv", header=T)
R = round(cor(rdataPCA2),digits=2)
myEigen = eigen(R)
rdataPCA2.pc = prcomp(rdataPCA2, retx=TRUE, center=TRUE,scale.=TRUE)
sd2 <- myEigen$values
loadings <- rdataPCA2.pc$rotation
rownames(loadings) <- colnames(rdataPCA2)
scores <- rdataPCA2.pc$x

# scree plot
plot(sd2,xlab="component number",ylab="eigenvalue",pch=12,cex=1.3)

# this plots the eigenvectors
plot(scores[,1], scores[,2], xlab="PCA 1", ylab="PCA 2",type="n",xlim=c(-7,5),ylim=c(-5,5))
arrows(0,0,loadings[,1]*10,loadings[,2]*10, length=0.1,angle=20, col="red")
#1.2 is the measure of the position of text away from the arrow point
text(loadings[,1]*10*1.2,loadings[,2]*10*1.2,rownames(loadings), col="blue", cex=0.7)

loadings2 = read.csv("C:\\Users\\GEO.Madan\\Dropbox\\Rclass project\\loadings2.csv", header=T)
par(mfrow= c(2,2))
barplot(loadings2$PC1, main="barplot of component 1", xlab="component 1", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
```

barplot(loadings2$PC2, main="barplot of component 2", xlab="component 2", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
barplot(loadings2$PC3, main="barplot of component 3", xlab="component 3", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
barplot(loadings2$PC4, main="barplot of component 4", xlab="component 4", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)

#biplot is a built in function to plot both scores and loadings at the same time simultaneously
biplot(scores[,1:2], loadings[,1:2], xlab=rownames(scores),ylab=NULL, cex=0.4,xlim=c(-3,3),ylim=c(-3,3),var.axes=TRUE)

#PCA (rdataPCA3.csv)
rdataPCA3=read.csv("C:\\Users\\GEO.Madan\\Dropbox\\Rclass project\\rdataPCA3.csv", header=T)
R=round(cor(rdataPCA3),digits=2)
myEig =eigen(R)
rdataPCA3.pc= prcomp(rdataPCA3, retx=TRUE, center=TRUE,scale.=TRUE)
sd <- myEig$values
loadings <- rdataPCA3.pc$rotation
rownames(loadings) <- colnames(rdataPCA3)
scores <- rdataPCA3.pc$x

# scree plot
plot(sd,xlab="component number",ylab="eigenvalue",pch=12,cex=1.3)

# this plots the eigenvectors
plot(scores[,1], scores[,2], xlab="PCA 1", ylab="PCA 2",type="n", xlim=c(-7,5),ylim=c(-5,5))
arrows(0,0,loadings[,1]*10,loadings[,2]*10, length=0.1,angle=20, col="red")
#1.2 is the measure of the position of text away from the arrow point
text(loadings[,1]*10*1.2,loadings[,2]*10*1.2,rownames(loadings), col="blue", cex=0.7)

loadings3=read.csv("C:\\Users\\GEO.Madan\\Dropbox\\Rclass project\\loadings3.csv", header=T)
par(mfrow= c(2,2))
barplot(loadings3$PC1, main="barplot of component 1", xlab="component 1", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
barplot(loadings3$PC2, main="barplot of component 2", xlab="component 2", ylab="loadings", col=
c("blue","1","2","3","4","RED","green","pink","grey","cyan","magenta","orange"), names=
c("WTD","PS","GTh","WD","b","Aqty","c.lth","Q","s","T","k","W.u"), cex.names=0.4)
Cluster analysis (cluster analysis.csv)

- `clusterdata=read.csv("C:\Users\GEO.Madan\Dropbox\Rclass project\cluster analysis.csv", header=T)`
- `cluster=clusterdata[,,-c(1,2)]` # removing first two column from analysis
- `means= apply(cluster,2,mean)` #vector of means == in apply(), 2=columns, 1=rows -- 3rd arg can be any vector function
- `sds= apply(cluster,2,sd)` #vector of std deviations # could also use mad=mean absolute deviation
- `cluster.scale = scale(cluster,center=means,scale=sds)` #normalizes to z scores
- `cluster.dist = dist(cluster.scale)` #euclidean similarity
- `cluster.hclust = hclust(cluster.dist)` #clusters on the distance matrix
- `plot(cluster.hclust,main='cluster analysis of pumping test data')`
- `groups.4=cutree(cluster.hclust,4)`
- `table(groups.4)`
- `sapply(unique(groups.4),function(g) clustdata$samples[groups.4==g])` #showing the individual data points in each group with sample number as its identification.
- `table(groups.4,clustdata$Township)` # representing a matrix of groups vs townships of the dataset.
- `aggregate(clustdata[,,-c(1,2)],list(groups.4),median)` #exclude first two columns and find the median of rest of the data and distribute into the groups.