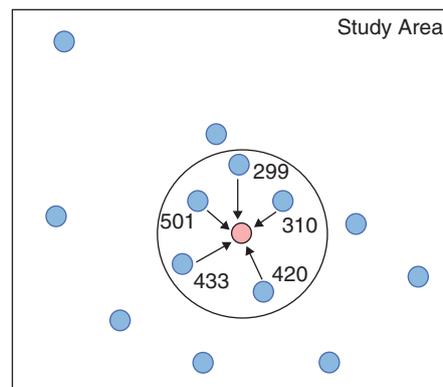


Introduction

Spatial Interpolation (SI) is a term used to estimate a value of a data variable at an un-sampled site from measurements made in close proximity or within a range of available data. This technique is based on Tobler's First Law of Geography that states that points close together in space are more likely to have similar values than points that are far apart. Use a neighborhood of sample points to estimate a value at an un-sampled location (figure below).



Interpolation uses a neighborhood of sample points of known values (blue color) to estimate a value at an un-sampled location (rose color). This method of estimation is using a specific radius from the un-sampled point.

Various interpolation techniques are used, and these techniques use sample values and X, Y coordinates to estimate the value of an un-sampled point. In general, different methods will generate unlike results with the same input data and no method is more accurate than others under all conditions. Users seeking accuracy should take in consideration a number of point samples and knowledge of the study area.

In order to produce a continuous representation of the phenomenon in question, interpolation makes use of sampling data. There are various methods in GIS that use the interpolation method. Deterministic interpolation techniques create surfaces from measured points. They are based on either the extent of the similarity, an example of such methods is the Inverse Distance Weighted (IDW). There is also a degree of smoothing such as the Trend Surface Analysis method. Geostatistical interpolation techniques such as Kriging are based on statistics and are used for more advanced prediction surface modeling, and also includes errors or uncertainty of predictions. Kriging method is based on the theory of regionalized variables and it is performed by placing an evenly spaced grid over the area for which we have known values and can obtain an estimated surface. The basic idea of Kriging interpolation is that every unknown point can be estimated by the weighted sum of the known points within a certain radius.

Electronic Supplementary Material: The online version of this chapter (https://doi.org/10.1007/978-3-319-61158-7_13) contains supplementary material, which is available to authorized users.

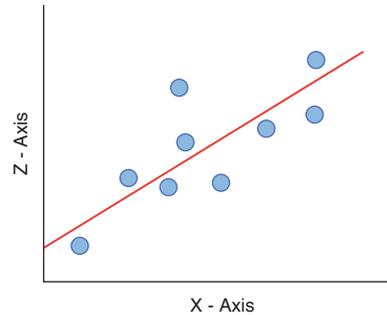
Method of Interpolation

Trend Surface Analysis

Trend surface analysis is a simple way for describing large variations and its function is to find general tendencies of the sample data, rather than to model a surface precisely. The trend analysis calculates the coefficients of a best-fit polynomial surface to fit a set of spatially distributed data points.

In one dimension (1-D): z varies as a linear function of x

$$Z = b_0 + b_1x + e$$



In two dimensions (2-D): z varies as a linear function of x and y

$$Z = b_0 + b_1x + b_2y + e$$

where Z is the interpolated parameter.

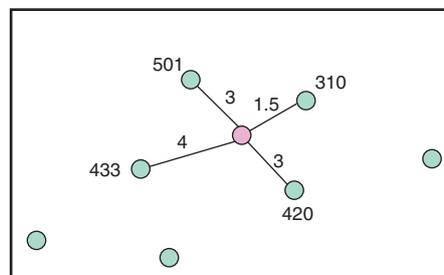
X and Y are the coordinates of the wells.

b coefficient is estimated from the control points. e : error in prediction.

The aim of this method is to develop a general kind of spatial distribution of an observable fact. The surface can be modeled using a linear or trend surface. Linear trends describe only the major direction and rate of change, while the trend surface provides progressively more complex descriptions of spatial patterns.

Inverse Distance Weighting (IDW)

Inverse distance weighting is a very popular technique in GIS and considers one of the simplest interpolation methods. There are a variety of methods that use weighted moving averages of points within a zone of influence. Interpolation techniques in which interpolation estimates are made based on values at nearby locations weighted only by distance from the interpolation location (figure below).



IDW: Closest 4 neighbors

In general the simplified formula for IDW is:

$$V_0 = \frac{\sum_{i=1}^n \left(\frac{V_i}{D_i} \right)}{\sum_{i=1}^n \left(\frac{1}{D_i} \right)}$$

where V_0 is the predictable value at point 0, V_i is the V value at control point i, D_i is the distance between control point i and 0, and n is the number of known values used in the evaluation.

The weights are a decreasing function of distance and the user has control over the mathematical form of the weighting function. The size of the neighborhood can be expressed as a radius or a number of points.

Global Polynomial (GP)

Global Polynomial or GP fits a smooth surface that is defined by a polynomial to the input sample points such as the TDS field in the attribute table of the well layer. The GP is similar to taking a piece of paper and fitting it in between the raised TDS values. The result from GP interpolation is a smooth surface that represents gradual trends in the surface over the area of interest. It is used by fitting a surface to the sample points when the surface varies slowly from region to region over the area of interest. While examining and/or removing the effects of long-range or global trends. In such circumstances, the technique is often referred to as trend surface analysis.

Kriging

Using geostatistical techniques, you can create surfaces incorporating the statistical properties of the measured data. Kriging is based on statistics. These techniques produce not only prediction surfaces but also error or uncertainty surfaces, giving you an indication of how good the predictions are. Many kriging methods are associated with geostatistics, but they are all in the kriging family. Ordinary, simple, universal, probability, indicator, and disjunctive kriging, along with their counterparts in cokriging, are all available in the Geostatistical Analyst. Not only do these kriging methods create predictions and error surfaces, they can also produce probability and quantile output maps depending on user needs. Kriging is the estimation procedure using known values and a semi-variogram to determine unknown values. The procedures involved in kriging incorporate measures of error and uncertainty when determine estimations. Based on the semi-variogram used, optimal weights are assigned to unknown values in order to calculate the unknown ones. Since the variogram changes with distance, the weights depend on the known sample distribution. The basic equation used in ordinary kriging is as follows:

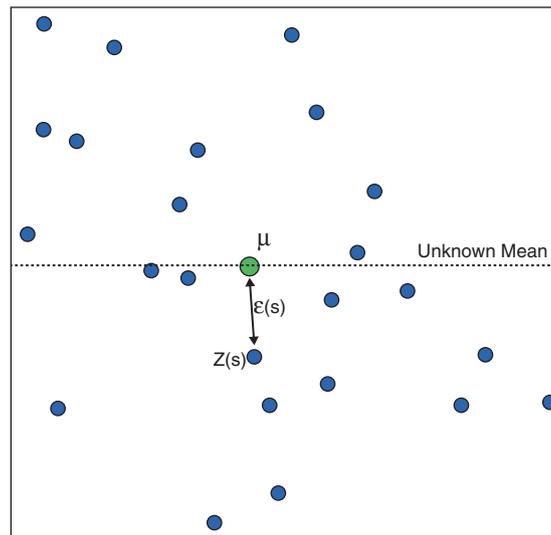
$$K(d) = \frac{1}{2n} \sum_{i=1}^n (Z(x_i) - Z(x_i + d))^2$$

where d is the distance between known points, n is the number of pairs of sample separated by d; Z is the attribute value (elevation of known points). The equation indicates that the semi-variance is expected to increase as d increases.

One of the most popular approaches is the ordinary kriging, which will be applied in this study. Ordinary kriging assumes the model:

$$Z(\mathbf{s}) = \mu + \varepsilon(\mathbf{s}),$$

where μ is an unknown constant. One of the main issues concerning ordinary kriging is whether the assumption of a constant mean is reasonable. Sometimes there are good scientific reasons to reject this assumption. However, as a simple prediction method, it has remarkable flexibility. The following figure is an example in one spatial dimension:



The data is a well with TDS values collected from Maawil watershed in Oman. The wells' locations look like they are distributed randomly. The data is simulated from the ordinary kriging model with a constant mean μ . The true but unknown mean is given by the dashed line. Thus, ordinary kriging can be used for data that seems to have a trend. There is no way to decide, based on the data alone, whether the observed pattern is the result of autocorrelation among the errors $\epsilon(s)$ with μ constant or trend, with $\mu(s)$ changing with s .

Ordinary kriging can use either semi-variograms or co-variances (which are the mathematical forms you use to express autocorrelation), use transformations and remove trends, and allow for measurement error.

Scenario 1: You are a hydrogeologist working for the water resources in Oman. You have been given a task to evaluate the groundwater along the coast in the Maawil watershed. You have been asked to evaluate, first if the densities of the wells has an effect on the salt intrusion and second, if the quality of groundwater downstream of the two dams, Maawil and Al-Kabir has been improved. The two dams has been built to store the surface runoff produced by rain in the rainy season, and then the stored water in the two dams gradually infiltrate to recharge the aquifer and improve its water quality. To answer these two questions you have been asked to use a GIS technique.

Data and Coordinate System

There are four shapefiles **Dam.shp**, **Stream.shp**, **Watershed.shp** and **Well.shp**. The files are registered in UTM zone 40 and the datum is WGS 1972 (**WGS_1972_UTM_Zone_40N**). Maawil watershed has 1,758 groundwater wells drilled mainly in the upper Maawil catchment area (downstream from the Maawil and Al-Kabir dams). The wells contain complete information about salinity (TDS) and nitrate (NO_3).

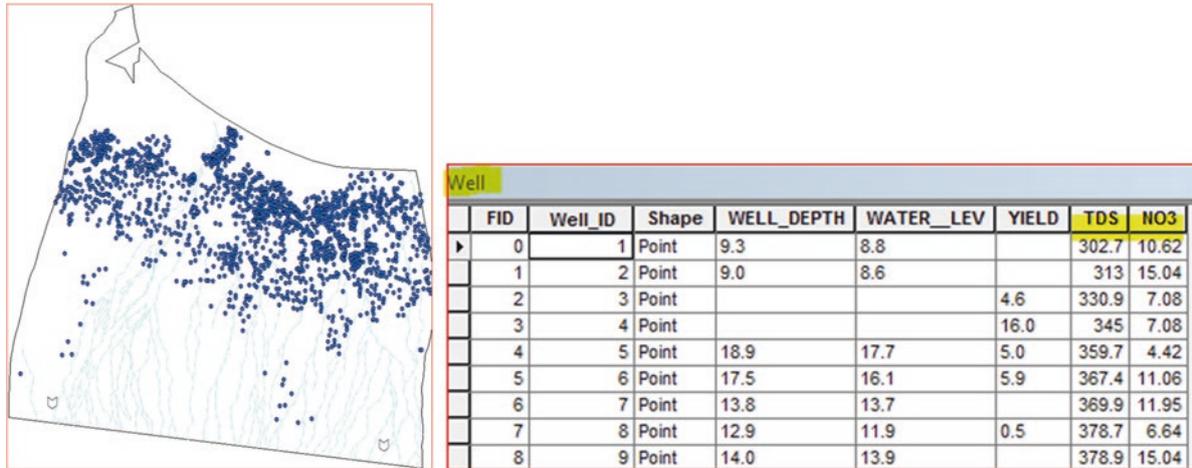
Density of Groundwater Well

The Point Density calculates a magnitude-per-unit area from point features that fall within a neighborhood around each cell. Adopting a bigger radii yields a more generalized density raster, and a smaller radius yields a more detailed raster. Only the wells that fall within the neighborhood are considered when calculating the density. If no wells fall within the neighborhood at a particular cell, that cell is assigned no data (NoData).

GIS Approach

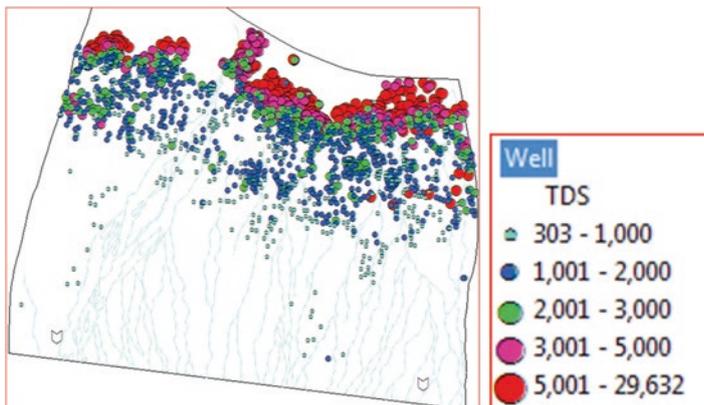
1. Start ArcMap and integrate the following file: **Dam.shp, Stream.shp, Watershed.shp** and **Well.shp** from \\Data\Q1 folder
2. Rename the Layers data frame “Well Density”
3. Symbolize the Dam, Stream, Watershed and Well layers using proper symbols
4. R-click Well.shp/Open Attribute Table

Result: The table contains two fields (TDS and NO₃) that are subject to analysis, close the well table



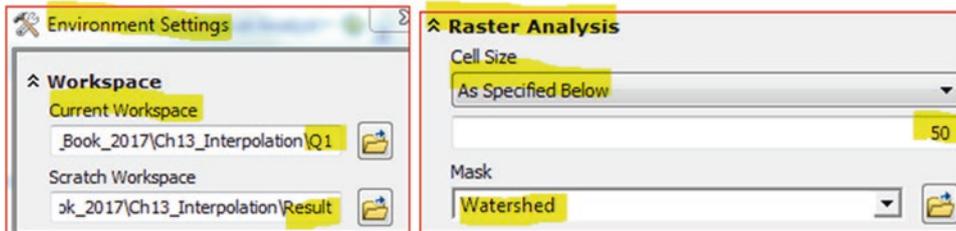
Symbolize the Wells Based on TDS

5. Double click Well layer/Symbology/Quantities/Graduate Symbols/Field Value = TDS/Classes 5/click Classify/ Method = Manual/Under Break Values, type 1000, 2000, 3000, 5000, and leave the last number (29632) as it is/click an empty place and click OK
6. Click Label header/Format Labels/Decimal = 0/Check Show thousand separator/OK
7. R-click each symbol of the Well layer in the TOC, change the color based on your taste



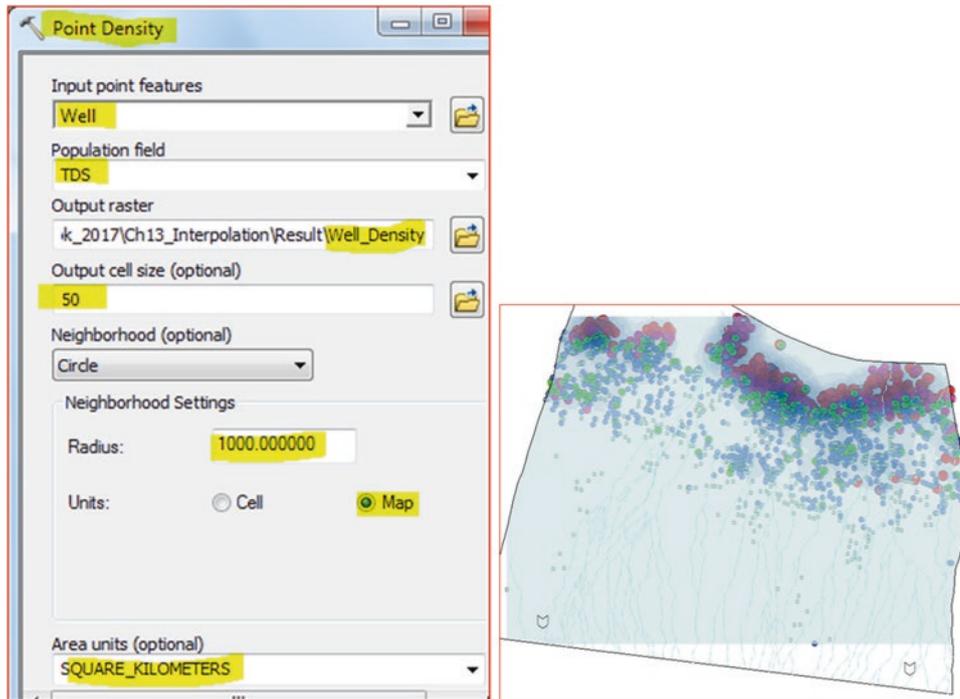
Result: The highest salinity with the bigger symbols is displayed along the coast of Oman.

8. Launch ArcToolbox
9. Right click an empty place below ArcToolbox/Environment/click Workspace
 - a. Current Workspace: \\Data\Q1
 - b. Scratch Workspace: \\Result
10. Open Raster Analysis in the Environment Setting
 - a. Cell Size: As Specified Below
 - b. Type: 50
 - c. Mask Watershed
11. Click OK



12. ArcToolbox/Spatial Analyst Tools/Density
13. Double click Point Density
14. Input point features: **Well**
15. Population field: **None**
16. Output raster: \\Result\Well_Density.tif
17. Output cell: 50
18. Neighborhood Circle
19. Radius: 50
20. Units: Cell
21. Area Units: SQUARE_KILOMETERS
22. Click OK
23. D-click Well_Density/Symbology/Label Header/Format Labels/Decimal = 0 and show thousands separators/OK/OK
24. Drag Well layer and place it below Well_Density
25. Right click Well_Density/Properties/Display tab/Transparency 30%/OK

Result: The Well_Density raster map displays and shows the cell densities using the salinity (TDS) field. Wells that have high salinity demonstrate bigger densities. This is true that the density of the wells in one location could affect the cone of depression through reducing the water table below the sea level of Gulf of Oman. This causes the sea saline water to invade the shallow aquifer along the coast and increase its salinity to a higher concentration.

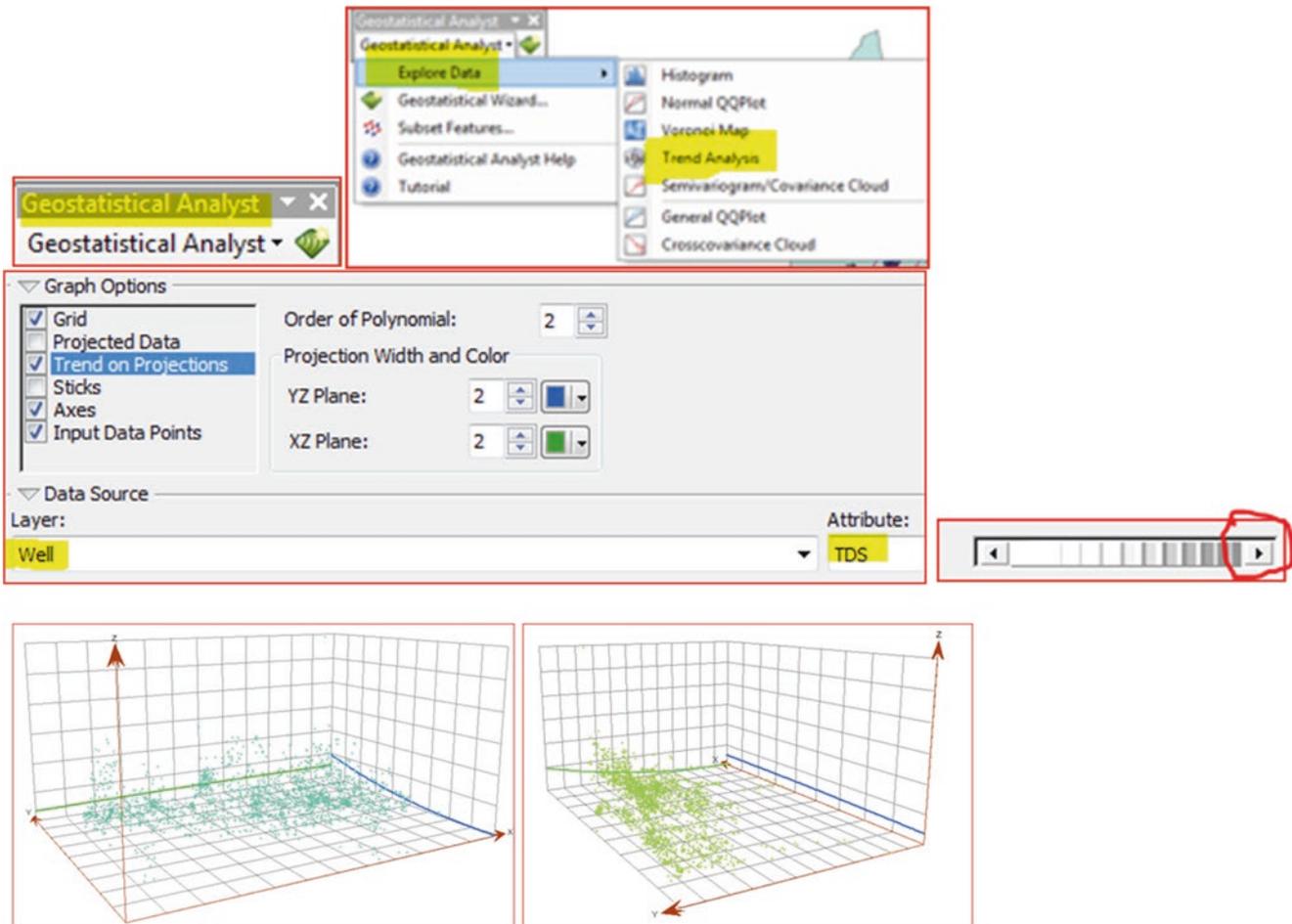


Trend Analysis

The Trend Analysis tool can help identify trends in the input dataset and provides a 3-D perspective of the data. The wells are plotted on the X,Y plane and TDS value is represented by the Z-dimension. Polynomials are then graphed on the scatterplots on the projected planes. An additional feature is that you can rotate the data to isolate directional trends. By default, the tool will select second-order polynomials to show trends in the data, but you may want to investigate polynomials of order one and three to assess how well they fit the data.

The trend surface analysis is a useful tool in early data analysis for delineating basic information and trends regarding the distribution of data. This type of analysis will be performed on the TDS to detect any trend in the salinity

26. Activate the Geostatistical Analyst (Customize/Toolbars)
27. Geostatistical Analyst/Explore Data/Trend Analysis
28. Uncheck Projected Data
29. Uncheck Sticks
30. Layer: Well
31. Attribute: TDS
32. Click the rotate arrow to rotate the graph



Result: The 3-D diagram shows two trend projections: The west-east trend (XZ – green line) and the north-south trend (YZ-blue). The XZ plane dips from west to east, which means that the TDS concentration increase toward east (toward the coast of Oman) and more or less in the north-south direction. This relationship is clear when you rotate the graph (second image).

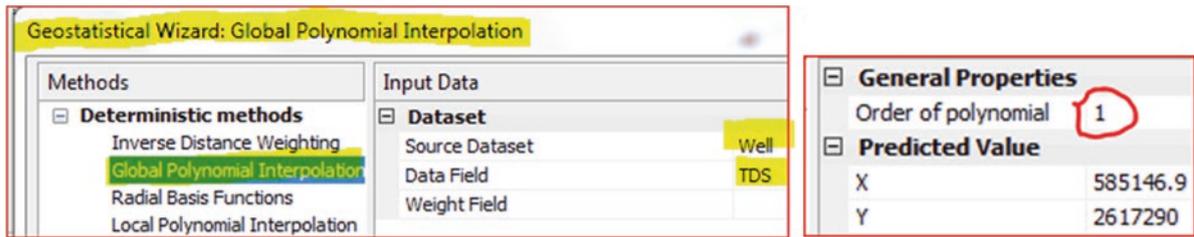
33. Close the Trend Analysis

I: Global Polynomial Interpolation

1. Geostatistical Analyst/Geostatistical Wizard
2. Methods
3. Deterministic methods: Global Polynomial Interpolation
4. Input Data/Dataset
5. Source Dataset: **Well**
6. Data Field: **TDS**
7. Next
8. Use Mean

Note: This step allows you to choose the order of polynomial from 1 to 10.

9. Select the **Power 1**
10. Next
11. The Root-Mean-Square (RMS) is **1936.07**
12. Repeat the process by choosing 2, 3, 4, 5, 6, and fill the table below

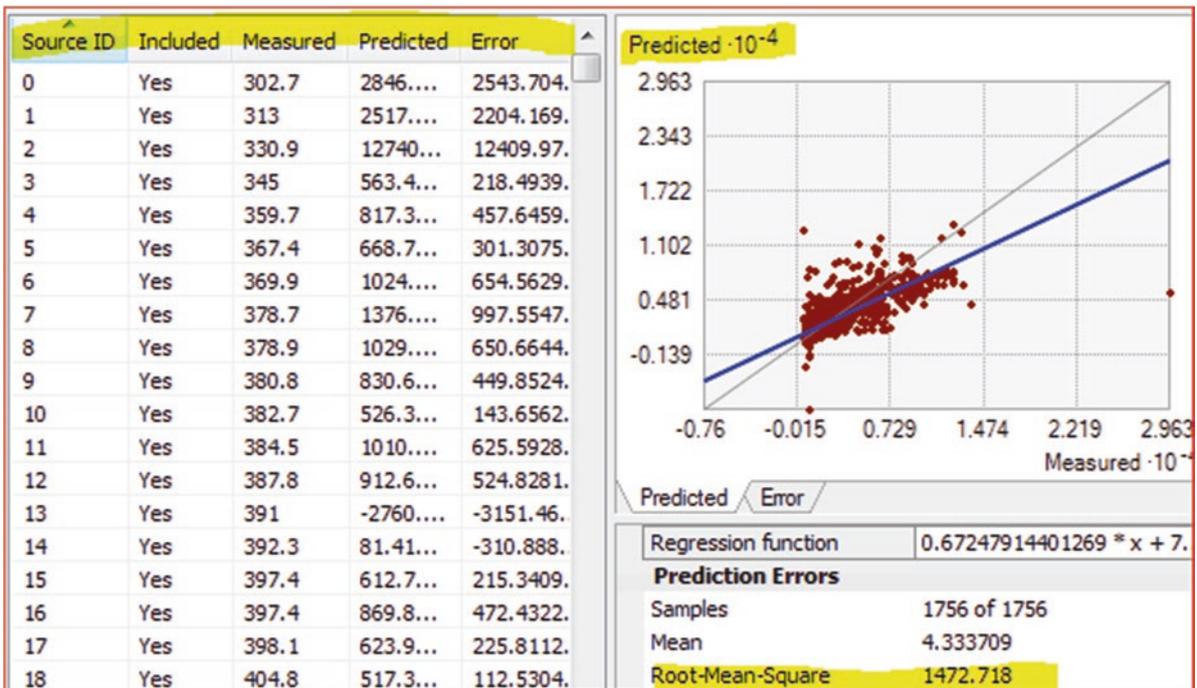


Power	RMS
1	1,936.07
2	1,819.97
3	1,772.62
4	1,727.58
5	1,686.02
6	1,741.25
7	1,777.44
8	1,472.72
9	271,845.00
10	20,758.40

Note: This step shows two things: a Scatter plot (predicted values vs measured values) and a Table (Measured, Predicted, and Error).

Result: The best polynomial is when the power is 8 as it generate lower RMS, which is 1472.72.

13. Go back and select the **Power 8**, then click Next
14. Click Finish then click OK

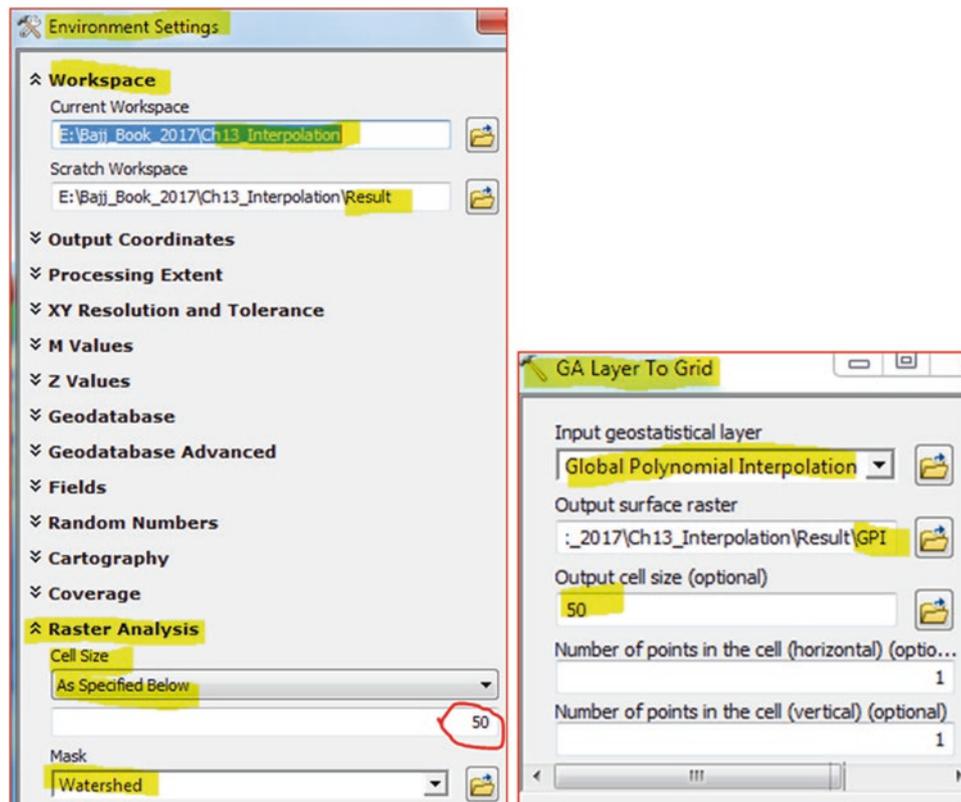


Result: Global Polynomial Interpolation Prediction Map layer is displayed in the TOC and has the same area extent of the Well layer.

Convert the GPI Predicted Map into ESRI Grid and Clip It Using the Mask Technique

Converting the Predicted Map output layer into ESRI grid is an essential step in order to clip the interpolated raster to fit the watershed area using the Mask in the Raster Analysis.

15. Right click GPI Prediction Map/Data/Export to Raster
16. Click the Environment Item and make sure the Workspace and Raster Analysis are set as below:
17. Open Workspace
 - a. Current Workspace: \\Data\Q1
 - b. Scratch Workspace: \\Result
18. Open Raster Analysis
 - a. Cell Size: As Specified Below
 - b. Type 50
 - c. Mask Watershed
19. OK
20. Input geostatistical layer: Global Polynomial Interpolation
21. Output Raster: \\Result\GPI
22. Output cell size: **50**
23. Click OK



Result: The GPI grid is created and added to the TOC and it is clipped to the Watershed area.

Classify the GPI Map

Classify the grid manually into 10 classes

24. Double click GPI/Symbology tab/under show: highlight Classified/Classes change to 10

25. Click Classify/Method = Manual
26. Under Break Values replace the numbers by 500, 1000, 1500, 2000, 3000, 4000, 5000, 10000, 15000, and 30000
27. Click OK
28. Click Label Header/Format Label/and make the decimal 0
29. Click OK
30. Change the first figure under label to “0–500”/OK
31. In the TOC, r-click the symbol of 15,001–30,000 and change the color into Tuscan Red.
32. Continue changing the color as in the table below

TDS Range	Color
332 - 500	Apatite Blue
501 - 1,000	Cretan Blue
1,001 - 1,500	Solar Yellow
1,501 - 2,000	Light Apple
2,001 - 3,000	Leaf Green
3,001 - 4,000	Fir Green
4,001 - 5,000	Lilac
5,001 - 10,000	Ginger Pink
10001 - 15,000	Mars Red
15,001 - 30,000	Tuscan Red

Result: The generated GPI salinity maps for the watershed area indicates that the interpolated surface downstream from the two dams was dominated by the low TDS with the exception from the lower left side of the watershed. In general, the water quality demonstrates improvement away from the coast of Oman.

II: Inverse Distance Weighting (IDW)

You will use now the Inverse Distance Weighting interpolation techniques

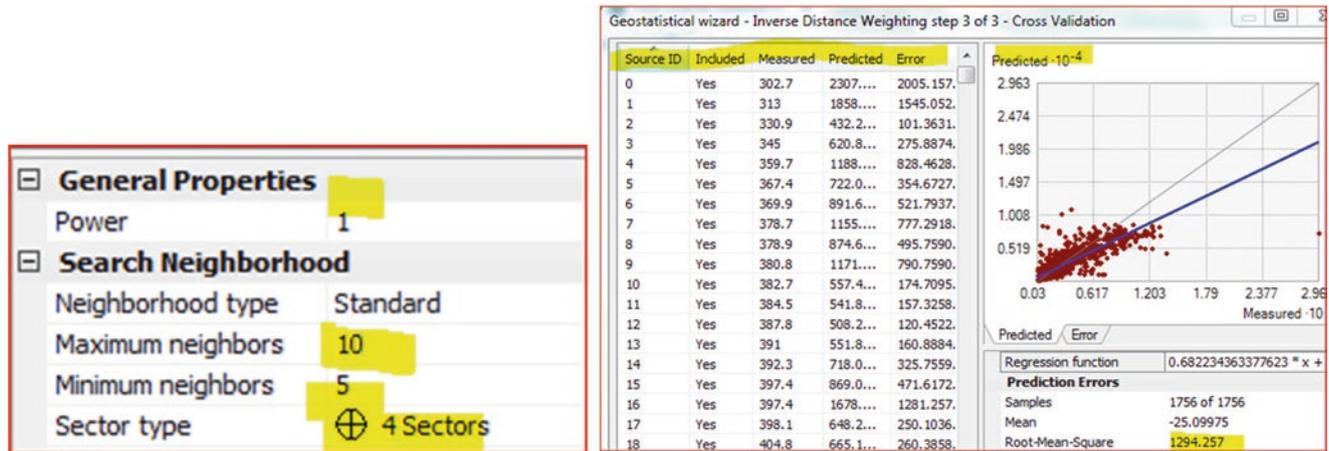
1. Insert Data Frame and call it IDW
2. Copy the Well, Stream, Dam, and Watershed layers into it from the GPI data frame
3. Geostatistical Analyst dropdown/Geostaistical Wizard
4. Deterministic methods: Inverse Distance Weighting
5. Source Dataset: Well
6. Data Field: TDS
7. Next
8. Check Use Mean
9. Click to optimize Power value  (icon in the upper right corner)

Result: The power changes from 2 to 1. This means that this is the optimal power value that will generate the minimum RMS, which is 1273.06.



Note: Geostatistical Analyst uses power values greater or equal to 1. When $p = 2$ (default value), the method is known as the inverse distance squared weighted interpolation. Although there is no theoretical justification to prefer this value over others. The effect of changing “p” should be investigated by previewing the output and examining the cross-validation statistics.

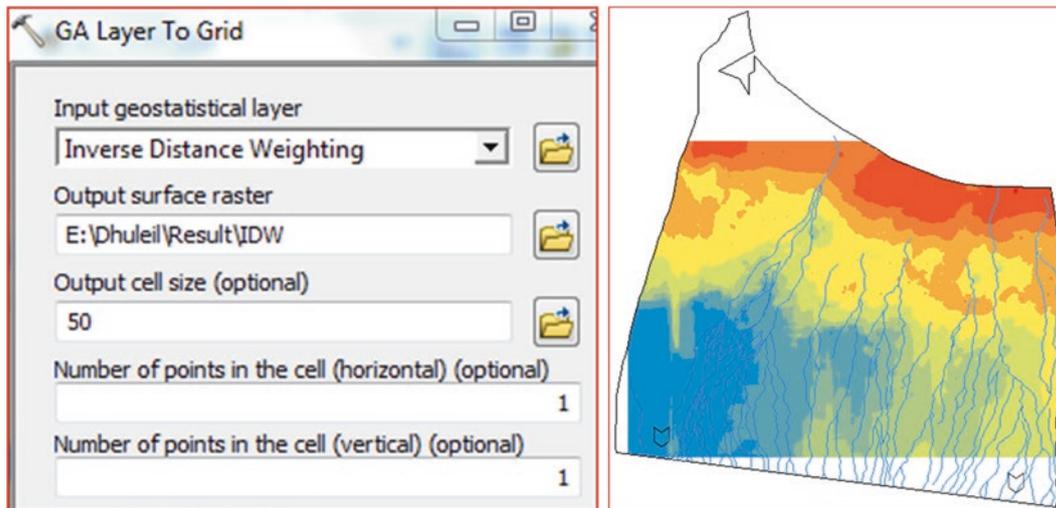
10. Change maximum and minimum neighborhood to 10 and 5 respectively
11. Sector type 4 sectors
12. Click next/Finish/Click OK



Result: Inverse Distance Weighting Prediction Map is an output layer and has the same area extend as the Well layer. The TDS ranges between 302.7 and 29,632 ppm.

Convert the IDW Predicted Map into ESRI Grid and Clip It Using the Mask Technique

13. Right click IDW Prediction Map/Data/Export to Raster
14. Click the Environment Item and make sure the Workspace and Raster Analysis are set as below:
15. Open Workspace
 - a. Current Workspace: \\Data\Q1
 - b. Scratch Workspace: \\Result
16. Open Raster Analysis
 - c. Cell Size: As Specified Below
 - d. Type 50
 - e. Mask Watershed
17. Click OK
18. Input geostatistical layer: Inverse Distance Weighting
19. Output Raster: \\Result\IDW
20. Output cell size: **50**
21. Click OK



Classify the IDW Map

Classify the **IDW** manually into 10 classes 500, 1000, 1500, 2000, 3000, 4000, 5000, 10000, 15000, and 30000

22. D-click IDW/Symbology tab/under show: highlight Classified/Classes 10
23. Click Classify/Method = Manual
24. Under Break Values replace and the numbers by 500, 1000, 1500, 2000, 3000, 4000, 5000, 10000, 15000, and 30000
25. Click OK
26. Click Label Header/Format Label/and make the decimal 0
27. Click OK
28. In the TOC, r-click the symbol of 15,001–30,000 and change the color into Tuscan Red.
29. Continue changing the color as in the table below

Break Values

500

1,000

1,500

2,000

3,000

4,000

5,000

10,000

15,000

29,279

TDS Range	Color
332 - 500	Apatite Blue
501 - 1,000	Cretan Blue
1,001 - 1,500	Solar Yellow
1,501 - 2,000	Light Apple
2,001 - 3,000	Leaf Green
3,001 - 4,000	Fir Green
4,001 - 5,000	Lilac
5,001 - 10,000	Ginger Pink
10001 - 15,000	Mars Red
15,001 - 30,000	Tuscan Red

Result: The generated IDW salinity maps for the watershed area indicates that the interpolated surface downstream the two dams was dominated by TDS less than 1,500 mg/l. The salinity of the wells are high along the coast and decrease in the west direction. This shows that the two dams are used as an artificial recharge and are improving the water quality of the aquifer.

Interpolation Using Kriging

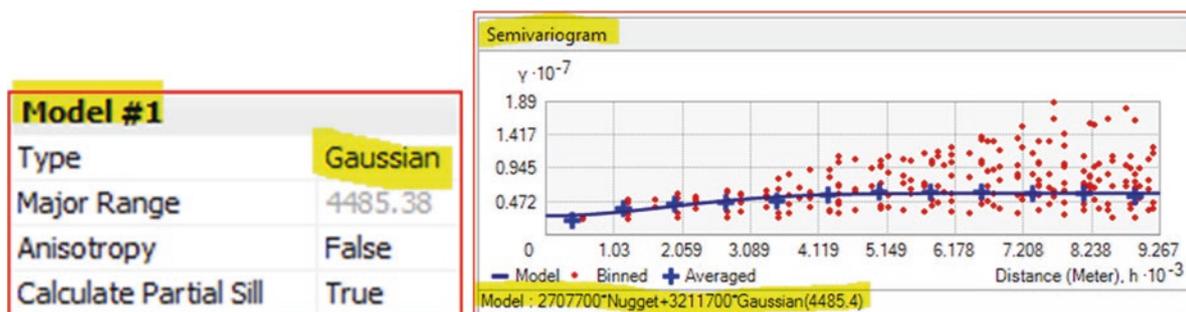
Kriging is the best possible interpolation method based on regression against observed z values of surrounding data points and weighted according to spatial covariance values. Kriging assigns weights according to a data-driven weighting function, rather than an arbitrary function. It is still just an interpolation algorithm and will give very similar results to others techniques such as IDW and GPI. There are various types of kriging and in this exercise, you will conduct an interpolation based on Ordinary Kriging.

30. Insert Data Frame and rename it Kriging
31. Copy the Well, Stream, Dam, and Watershed layers into it from the IDW data frame
32. Geostatistical Analyst dropdown/Geostatistical Wizard
33. Geostatistical methods: Kriging/CoKriging
34. Source Dataset: Well
35. Data Field: TDS
36. Next
37. Check Use Mean/OK
38. Kriging Type: Ordinary
39. Output Surface Type: Prediction
40. Next

Methods	Input Data
<ul style="list-style-type: none"> [-] Deterministic methods <ul style="list-style-type: none"> Inverse Distance Weighting Global Polynomial Interpolation Radial Basis Functions Local Polynomial Interpolation [-] Geostatistical methods <ul style="list-style-type: none"> Kriging / CoKriging 	<ul style="list-style-type: none"> [-] Dataset <ul style="list-style-type: none"> Source Dataset: Well Data Field: TDS [-] Dataset 2 <ul style="list-style-type: none"> Source Dataset: <none> [-] Dataset 3 <ul style="list-style-type: none"> Source Dataset: <none>

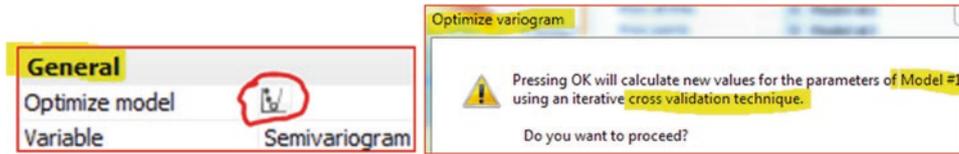
Geostatistical wizard - Kriging step 2 of 5	
Kriging Type	<ul style="list-style-type: none"> Ordinary Simple Universal Indicator Probability Disjunctive
Output Surface Type	<ul style="list-style-type: none"> Prediction Quantile Probability Prediction Standard Error

41. Under Model # 1
42. Change the Type to "Gaussian"



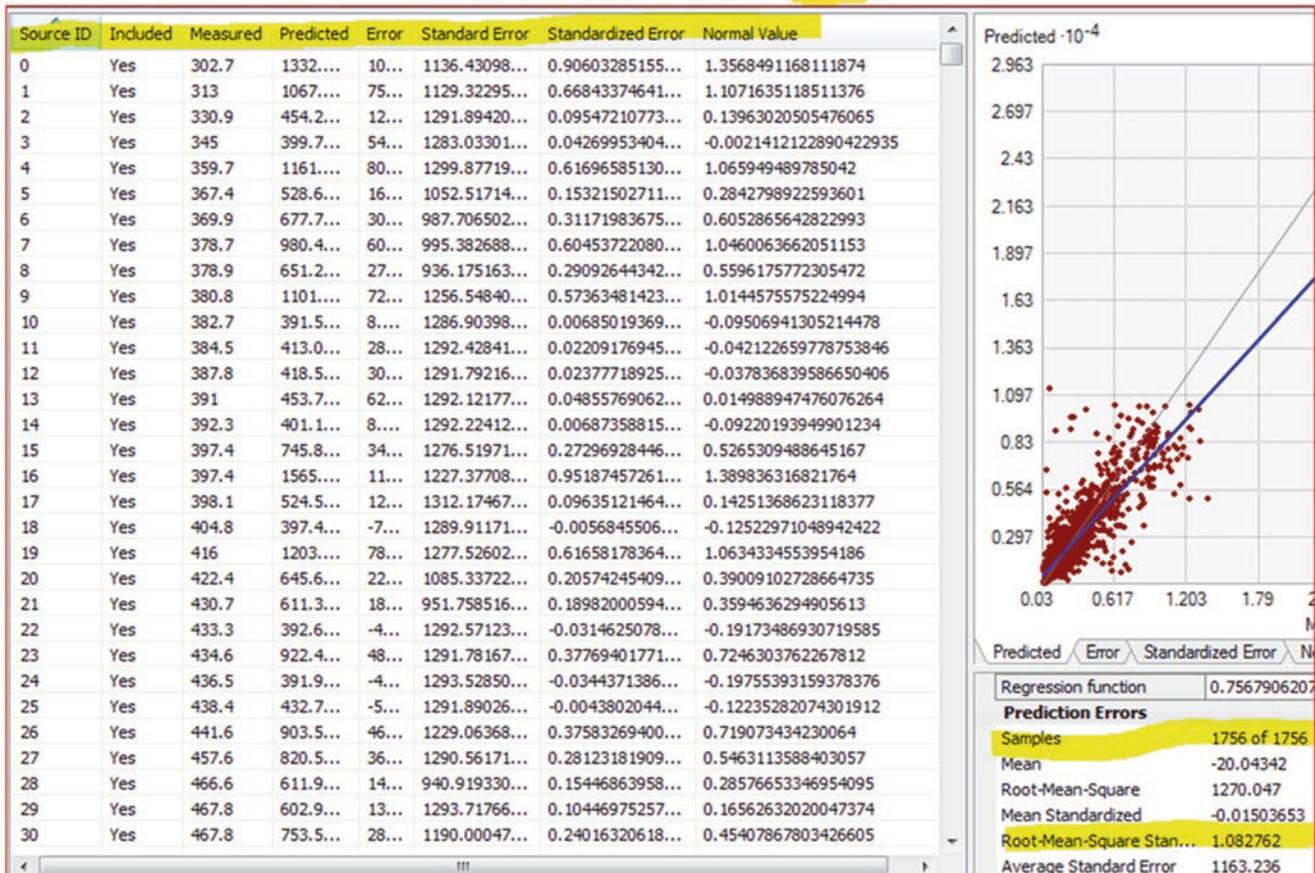
Note: The mathematical model that used in the Semi-variogram allows you to use different models (Stable, Circular, Spherical, Gaussian, Exponential, and others). In this exercise we are going to use the Gaussian model.

- 43. Under “General”
- 44. Click the Optimize model then click OK (to calculate the new value for the parameters of Gaussian model (Model # 1) using the iterative cross validation technique)



- 45. Click Next
- 46. Neighborhood Type: Standard
- 47. Maximum neighbors: 10
- 48. Minimum neighbors: 5
- 49. Sector Type 1 Sector
- 50. Next

Dataset	#0 [Well - TDS]
Search Neighborhood	
Neighborhood type	Standard
Maximum neighbors	10
Minimum neighbors	5
Sector type	<input checked="" type="radio"/> 1 Sector



- 51. Click Finish then click OK

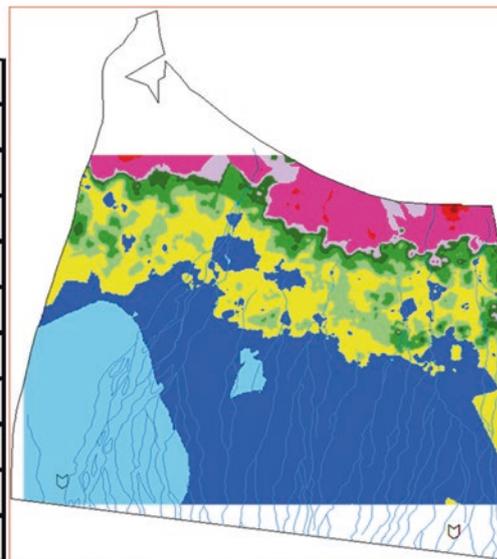
Result: Kriging Prediction Map is an output layer and has the same area extend as the Well layer. The TDS ranges between 302.7 and 29,632 ppm.

Convert the Kriging Predicted Map into ESRI Grid and Clip It Using the Mask Technique

52. Right click Kriging Prediction Map/Data/Export to Raster
53. Click the Environment Item
54. Open Workspace and make sure the Workspace and Raster Analysis are set as below:
 - f. Current Workspace: \\Data\Q1
 - g. Scratch Workspace: \\Result
55. Open Raster Analysis
 - h. Cell Size: As Specified Below
 - i. Type 50
 - j. Mask Watershed
56. Click OK
57. Input geostatistical layer: Kriging Prediction Map
58. Output Raster: \\Result\Kriging
59. Output cell size: **50**
60. Click OK
61. In Model # 1 Select the Type “Spherical”
62. Anisotropy select True/Next
63. Accept the Default/Next

Classify the **Kriging** manually into 10 classes 500, 1000, 1500, 2000, 3000, 4000, 5000, 10000, 15000, and 30000 as in the IDW and match the color as below.

TDS Range	Color
332 - 500	Apatite Blue
501 - 1,000	Cretan Blue
1,001 - 1,500	Solar Yellow
1,501 - 2,000	Light Apple
2,001 - 3,000	Leaf Green
3,001 - 4,000	Fir Green
4,001 - 5,000	Lilac
5,001 - 10,000	Ginger Pink
10001 - 15,000	Mars Red
15,001 - 30,000	Tuscan Red



Question1: Comment on the salinity interpolation.

Question2: Run the GPI, IDW, and Kriging using the nitrate (NO_3) variable.