

Geostatistical Analysis



ISTANBUL **TECHNICAL** UNIVERSITY

Sp. Anly. and Alg. in GIS
Week 10

Res. Assist. Ömer AKIN

Introduction & Aim of the Study

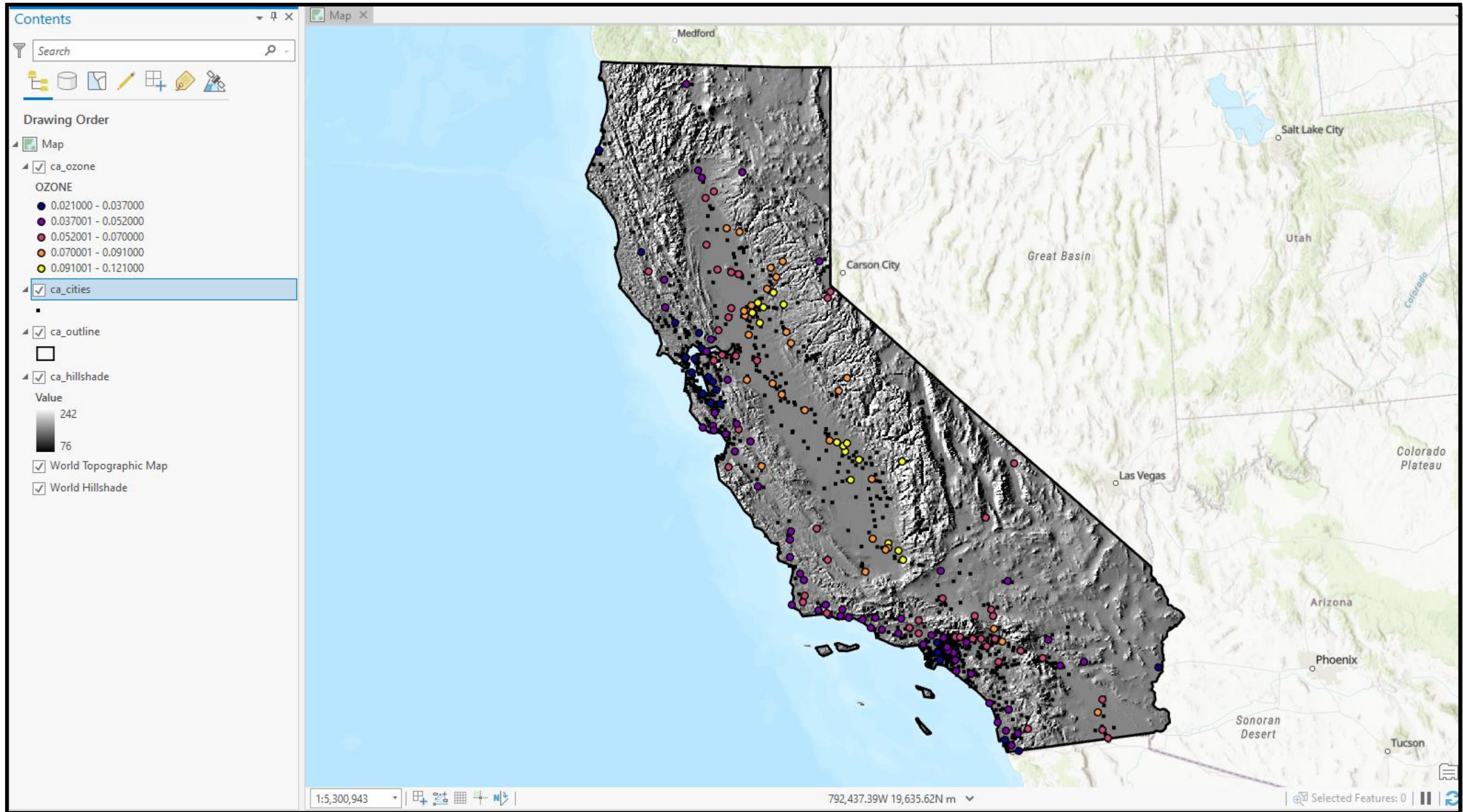
Aim of the Study:

- *Ozone concentration is measured at monitoring stations throughout the state of California. The locations of monitoring stations and concentration levels of ozone are known for all the stations, but the ozone values for other (unmonitored) locations in California are also of interest.*

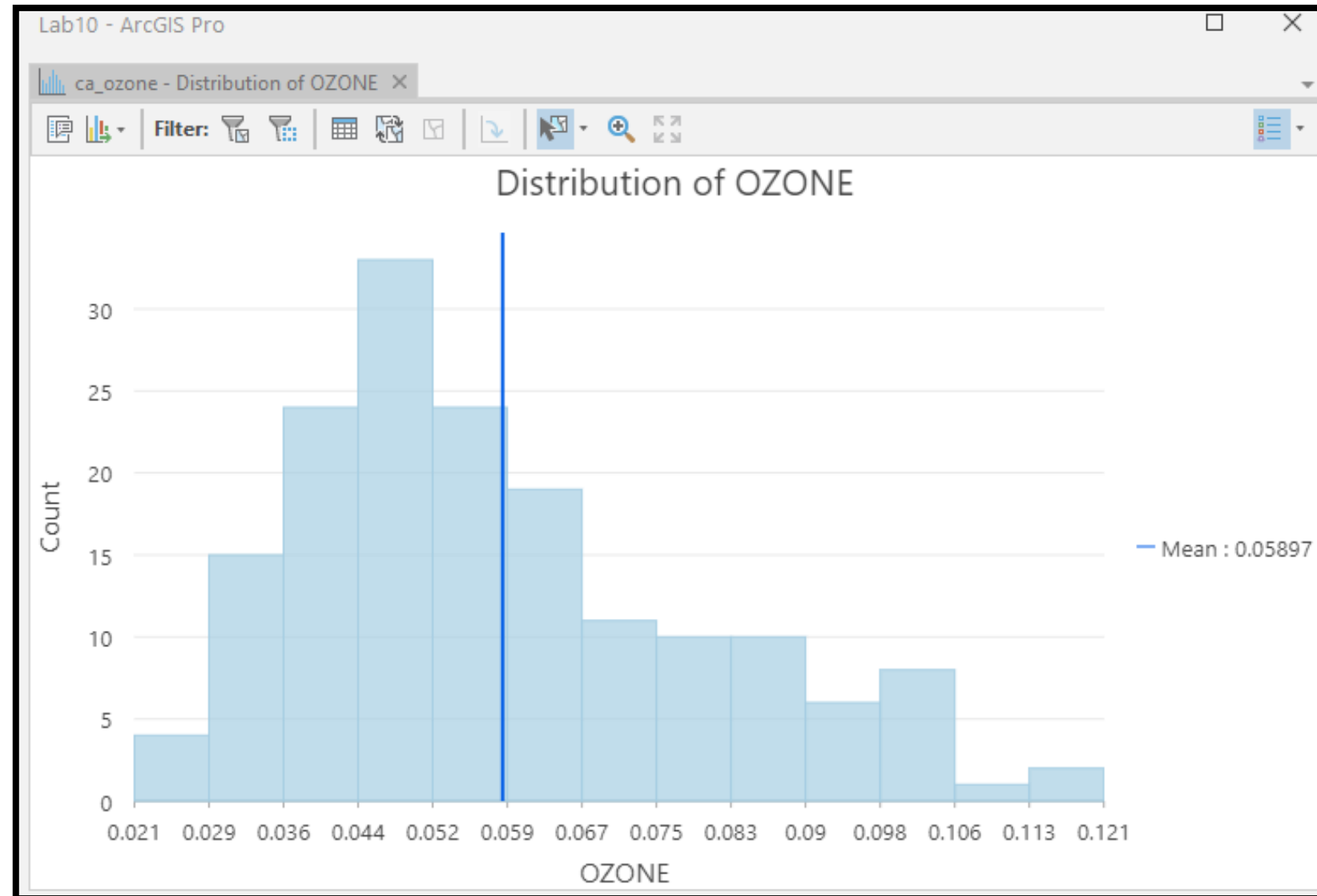
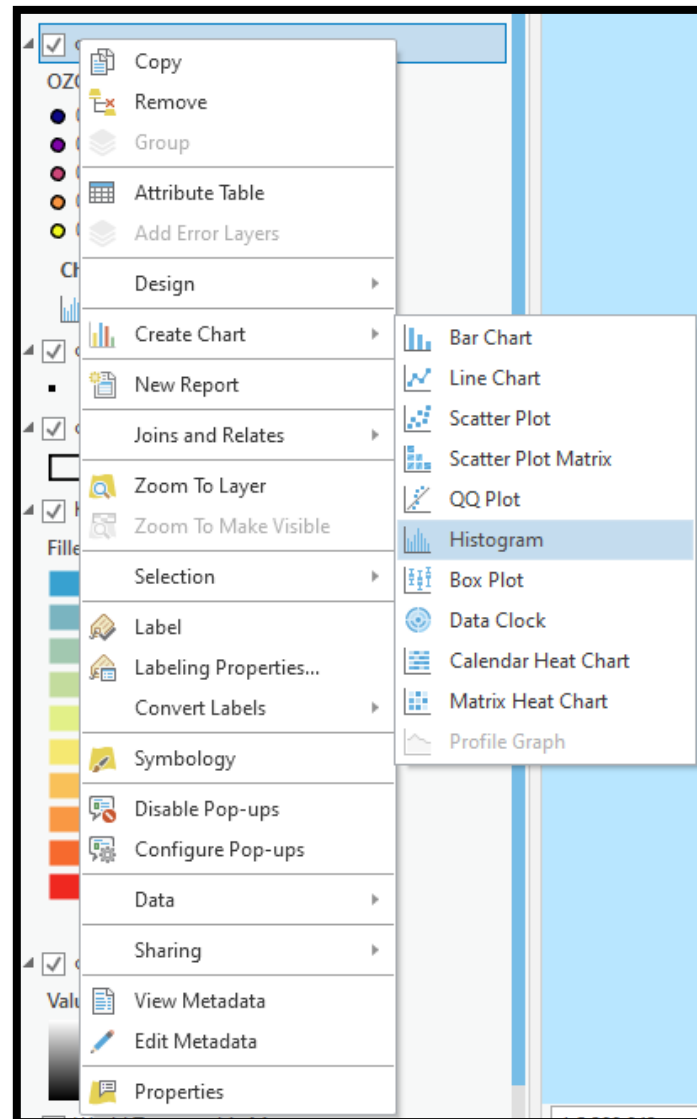
Input Data:

- *Ozone Concentration (measured on September 6, 2007 between 3:00 and 4:00 p.m.) (Vector-Point)*
- *City Centers (Vector-Point)*
- *Border of California (Vector-Polygon)*
- *Hillshade (Raster)*

Study Area & Data



Explore Data Statistics

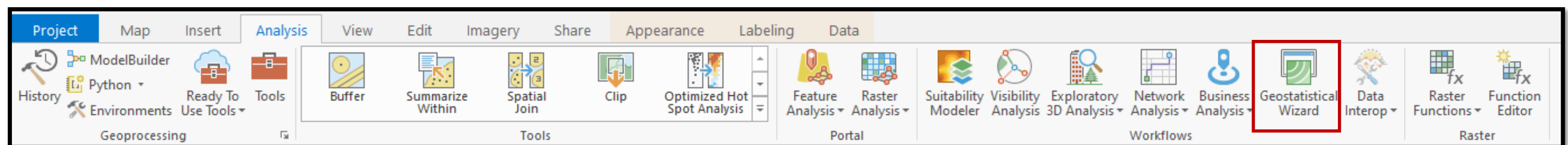


Statistics	
	Dataset
<input checked="" type="checkbox"/> Mean	0.0589730539
<input type="checkbox"/> Median	0.056
<input type="checkbox"/> Std. Dev.	0.0211534025
Count	167
Min	0.021
Max	0.121
Sum	9.8485
Nulls	0
Skewness	0.7042560409
Kurtosis	2.8320571739

- As a quick check, if the mean and the median are approximately the same value, you have one piece of evidence that the data may be normally distributed.
- The ozone data histogram indicates that the data is unimodal (one hump) and skewed right. The right tail of the distribution indicates the presence of a relatively small number of sample points with large ozone concentration values. It seems that the data is not close to a normal distribution.

Geostatistical Analyst

Geostatistical Analyst provides tools that make optimal predictions possible by examining the relationships between all the sample points and producing a continuous surface of ozone concentration, standard errors (uncertainty) of predictions, and probabilities that critical values are exceeded.



Geostatistical methods

☐ Empirical Bayesian Kriging
 ☐ EBK Regression Prediction
 ☒ **Kriging / CoKriging**
☐ Areal Interpolation

3D Interpolation

☐ Empirical Bayesian Kriging 3D

Interpolation with barriers

☐ Kernel Interpolation
 ☐ Diffusion Interpolation

Deterministic methods

☐ Local Polynomial Interpolation
 ☐ Inverse Distance Weighting
 ☐ Radial Basis Functions
 ☐ Global Polynomial Interpolation

Input Dataset 1

Source Dataset: cross_valatation_result

Data Field: Measured

Input Dataset 2

Source Dataset:

Data Field:

Kriging / CoKriging

Kriging is the oldest and most studied geostatistical interpolation method. It is very flexible and allows you to investigate graphs of spatial auto- and cross-correlation. Kriging uses statistical models that allow a variety of output surfaces including predictions, prediction standard errors, probability, and quantile. The flexibility of kriging can require a lot of decision-making. Kriging assumes the data come from a stationary stochastic process, and some methods assume normally-distributed data.

[Learn more about how Kriging works](#)

Interpolation Methods

Brief Information about technique

Detailed Information links

Kriging Methods / Ordinary Kriging

Geostatistical Wizard - Kriging / CoKriging

Geostatistical methods

- ☐ Empirical Bayesian Kriging
- ☐ EBK Regression Prediction
- ☒ **Kriging / CoKriging**
- ☐ Areal Interpolation

3D Interpolation

- ☐ Empirical Bayesian Kriging 3D

Interpolation with barriers

- ☐ Kernel Interpolation
- ☐ Diffusion Interpolation

Deterministic methods

- ☐ Local Polynomial Interpolation
- ☐ Inverse Distance Weighting
- ☐ Radial Basis Functions

Input Dataset 1

Source Dataset: ca_ozone

Data Field: OZONE

Input Dataset 2

Source Dataset:

Data Field:

Kriging / CoKriging

Kriging is the oldest and most studied geostatistical interpolation method. It is very flexible and allows you to investigate graphs of spatial auto- and cross-correlation. Kriging uses statistical models that allow a variety of output surfaces including predictions, prediction standard errors, probability, and quantile. The flexibility of kriging can require a lot of decision-making. Kriging assumes the data come from a stationary stochastic process, and some methods assume normally-distributed data.

[Learn more about how Kriging works](#)

< Back Next > Finish

Geostatistical Wizard - Kriging

Ordinary Kriging

- ☒ **Prediction**
- ☐ Quantile
- ☐ Probability
- ☐ Prediction Standard Error

Simple Kriging

- ☐ Prediction
- ☐ Quantile
- ☐ Probability
- ☐ Prediction Standard Error

Universal Kriging

- ☐ Prediction
- ☐ Quantile
- ☐ Probability
- ☐ Prediction Standard Error

Indicator Kriging

- ☐ Probability
- ☐ Standard Error of Indicators

Probability Kriging

- ☐ Probability
- ☐ Standard Error of Indicators

Disjunctive Kriging

Dataset #1

Transformation type: None

Order of Trend Removal: None

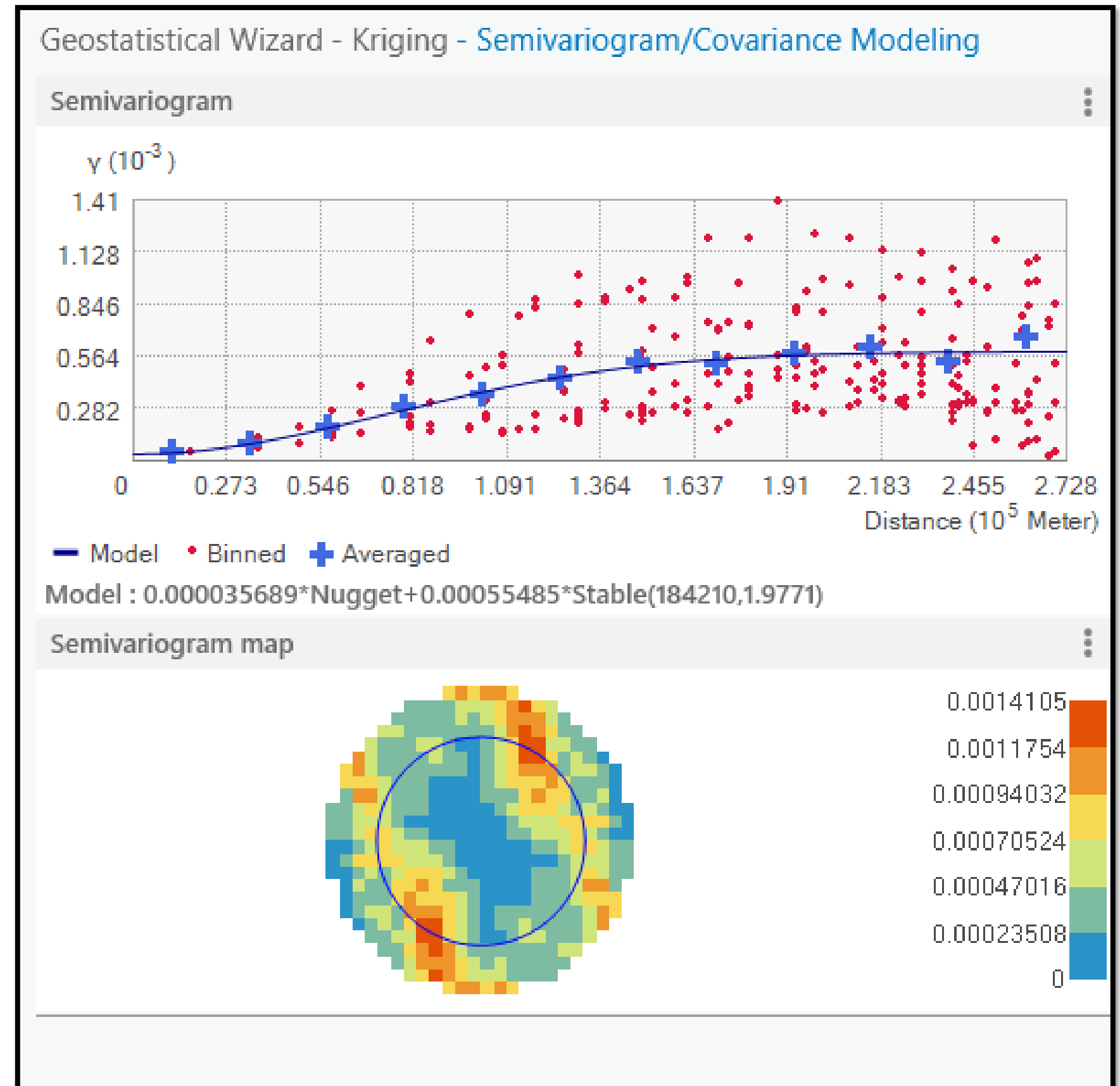
< Back Next > Finish

*To get more information about kriging techniques please visit:

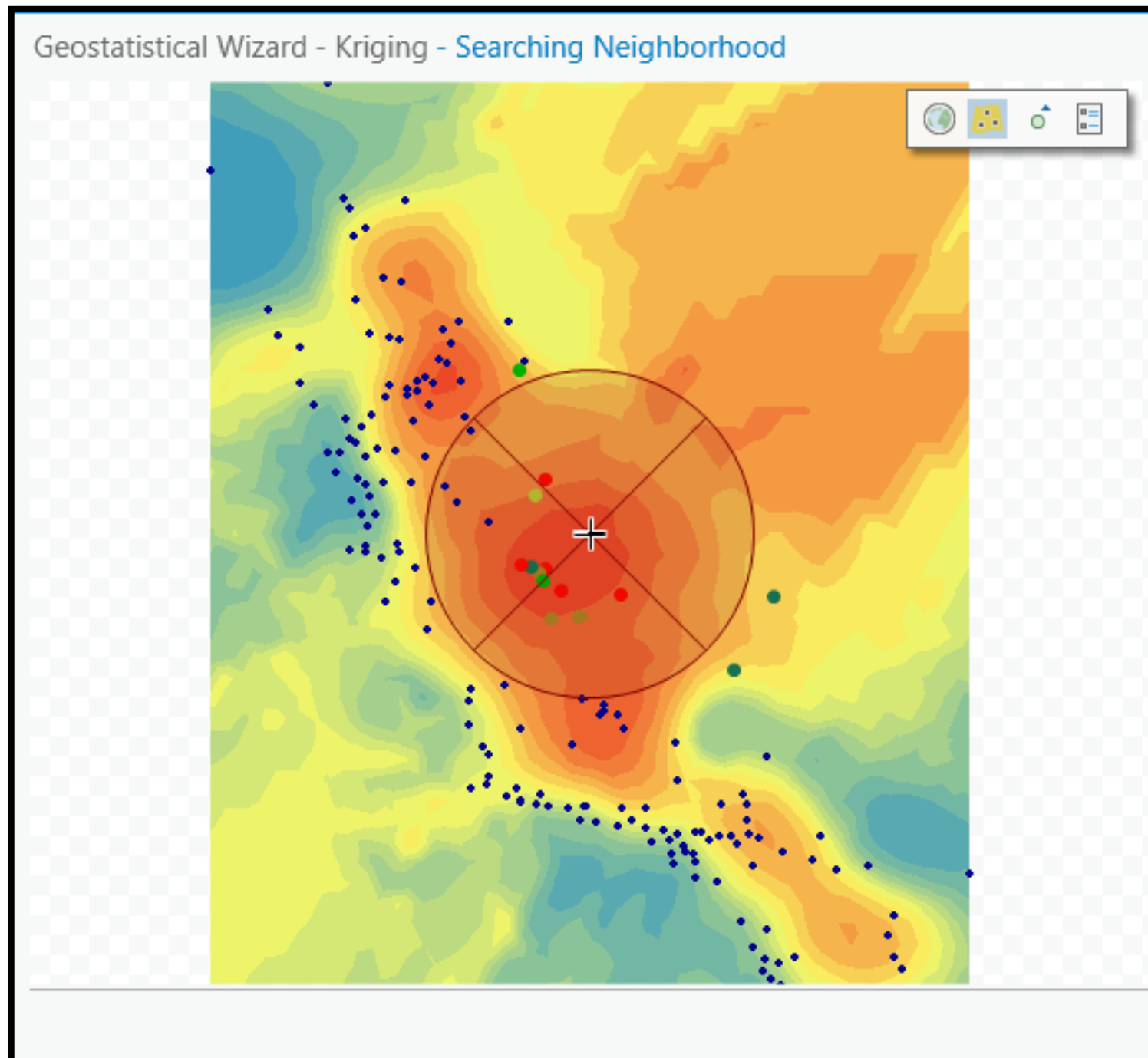
<https://pro.arcgis.com/en/pro-app/2.7/help/analysis/geostatistical-analyst/what-are-the-different-kriging-models-.htm>

Semivariogram/Covariance Model

- The semivariogram/covariance model is displayed, allowing you to examine spatial relationships between measured points.
- You can assume that things that are closer together are more alike than things that are farther apart. The semivariogram allows you to explore this assumption.
- The process of fitting a semivariogram model to capture the spatial relationships in the data is known as variography.

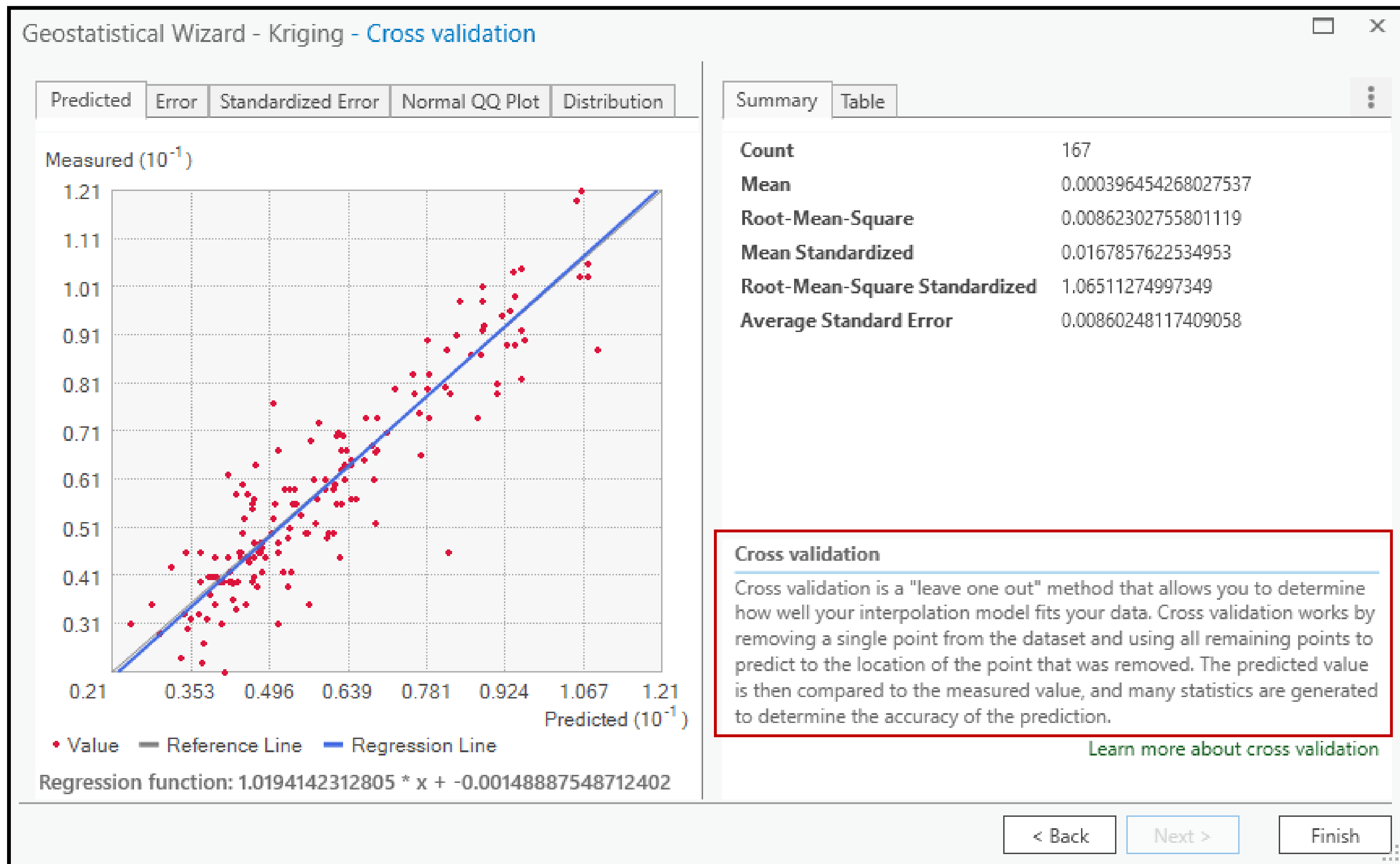


Prediction Surface

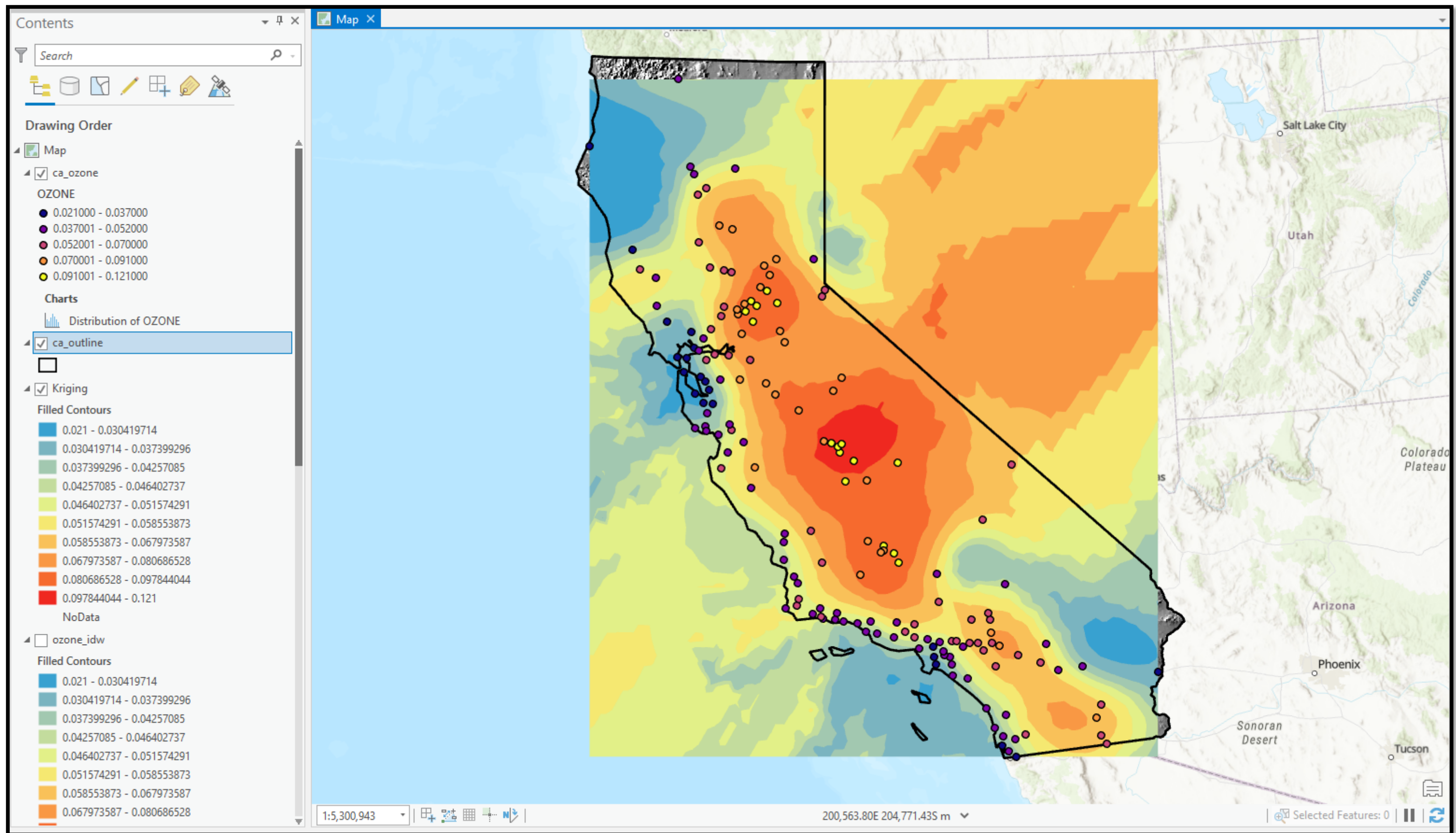


- As shown, the values of the closest measured locations are most alike to the value of the unmeasured location that you are trying to predict.
- The red points in the image below are going to be weighted (or influence the unknown value) more than the green points since they are closer to the location you are predicting.
- Using the surrounding points and the semivariogram/covariance model fitted previously, you can predict values for the unmeasured location.

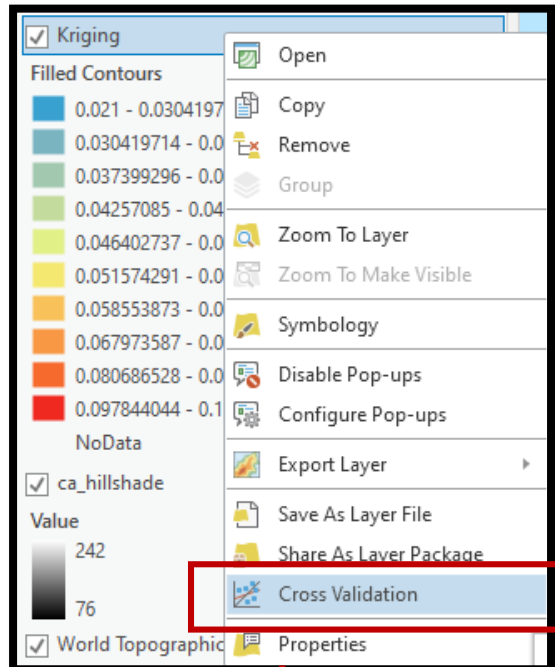
Cross Validation



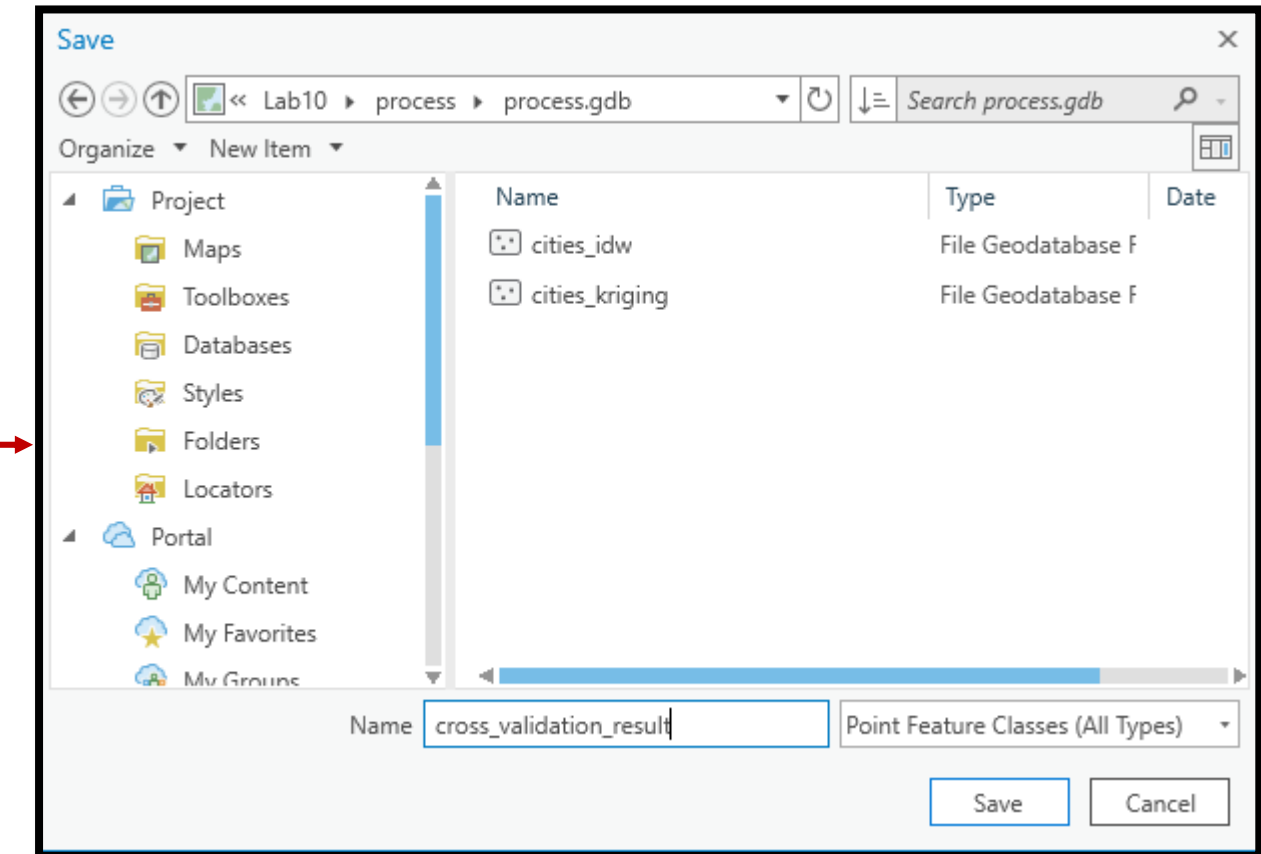
Result of Ordinary Kriging



Get Cross Validation Results as Feature



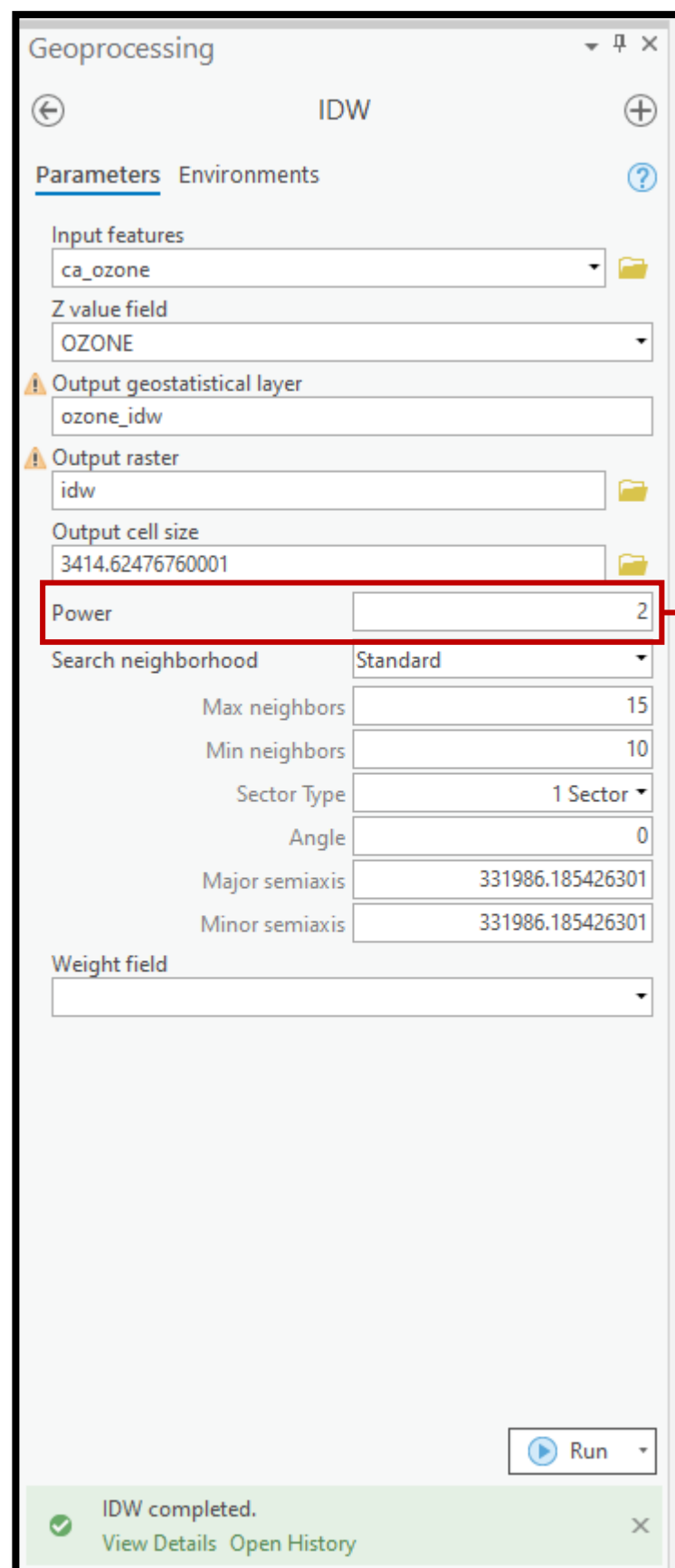
*Export table
as a feature*



Source ID	Included	Measured	Predicted	Error	StdError	StdDev	NormValue	Source_ID	Included
1	Yes	0.045	0.04864671218976611	0.003646712189766109	0.006731	0.541797	0.552161	1	Yes
2	Yes	0.106	0.1073464183423576	0.0013464183423576026	0.006655	0.202314	0.165865	2	Yes
3	Yes	0.04	0.04114055929533589	0.001140559295335887	0.006654	0.171417	0.120368	3	Yes
4	Yes	0.041	0.03930320405125129	-0.0016967959487487125	0.007052	-0.240605	-0.352383	4	Yes
5	Yes	0.09	0.09581881579045078	0.005818815790450779	0.008437	0.689688	0.717519	5	Yes
6	Yes	0.053	0.044778450227717516	-0.008221549772282483	0.009722	-0.845652	-0.882983	6	Yes
7	Yes	0.031	0.024132770906453133	-0.006867229093546867	0.00699	-0.982457	-1.024945	7	Yes
8	Yes	0.074	0.06916392909712589	-0.004836070902874107	0.006736	-0.717958	-0.756913	8	Yes
9	Yes	0.056	0.05383908660963144	-0.0021609133903685643	0.006725	-0.321327	-0.433474	9	Yes
10	Yes	0.066	0.0769167353221475	0.0109167353221475	0.007306	1.494121	1.399455	10	Yes
11	Yes	0.071	0.07078299356310337	-0.0002170064368966218	0.00753				
12	Yes	0.034	0.04338623966942537	0.009386239669425367	0.00785				
13	Yes	0.103	0.10741290809222281	0.00441290809222282	0.00715				
14	Yes	0.079	0.09092944989188251	0.011929449891882507	0.00645				
15	Yes	0.049	0.05979585875178074	0.01079585875178074	0.01217				
16	Yes	0.065	0.06672561933546507	0.00172561933546507	0.00677				
17	Yes	0.041	0.04655336141271305	0.005553361412713051	0.00651				
18	Yes	0.069	0.056980732271048266	-0.01201926772895174	0.00747				
19	Yes	0.064	0.06317824565362629	-0.0008217543463737154	0.00646				
20	Yes	0.067	0.0691040969649041	0.002104096964904101	0.00654				
21	Yes	0.061	0.05963335918542899	-0.0013666408145710102	0.00652				
22	Yes	0.04	0.043853073924053666	0.0038530739240536654	0.00637				
23	Yes	0.032	0.03501641986680172	0.00301641986680172	0.00656				
24	Yes	0.046	0.04789374698761509	0.0018937469876150886	0.00780				

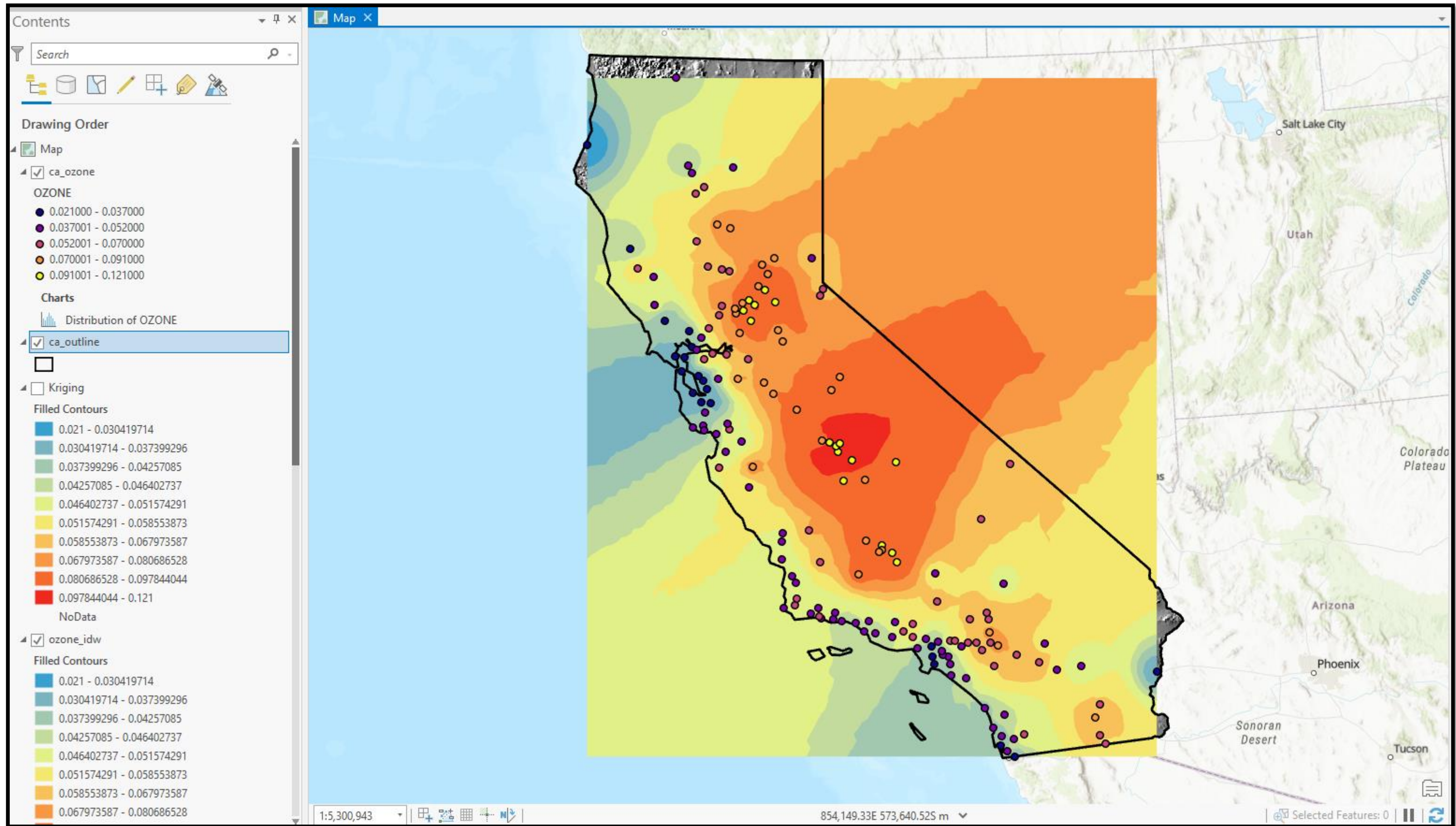
FID	Shape	Measured	Predicted	Error	StdError	StdDev	NormValue	Source_ID	Included
0	Point	0.045	0.048647	0.003647	0.006731	0.541797	0.552161	1	Yes
1	Point	0.106	0.107346	0.001346	0.006655	0.202314	0.165865	2	Yes
2	Point	0.04	0.041141	0.001141	0.006654	0.171417	0.120368	3	Yes
3	Point	0.041	0.039303	-0.001697	0.007052	-0.240605	-0.352383	4	Yes
4	Point	0.09	0.095819	0.005819	0.008437	0.689688	0.717519	5	Yes
5	Point	0.053	0.044778	-0.008222	0.009722	-0.845652	-0.882983	6	Yes
6	Point	0.031	0.024133	-0.006867	0.00699	-0.982457	-1.024945	7	Yes
7	Point	0.074	0.069164	-0.004836	0.006736	-0.717958	-0.756913	8	Yes
8	Point	0.056	0.053839	-0.002161	0.006725	-0.321327	-0.433474	9	Yes
9	Point	0.066	0.076917	0.010917	0.007306	1.494121	1.399455	10	Yes

Inverse Distance Weighted (IDW)

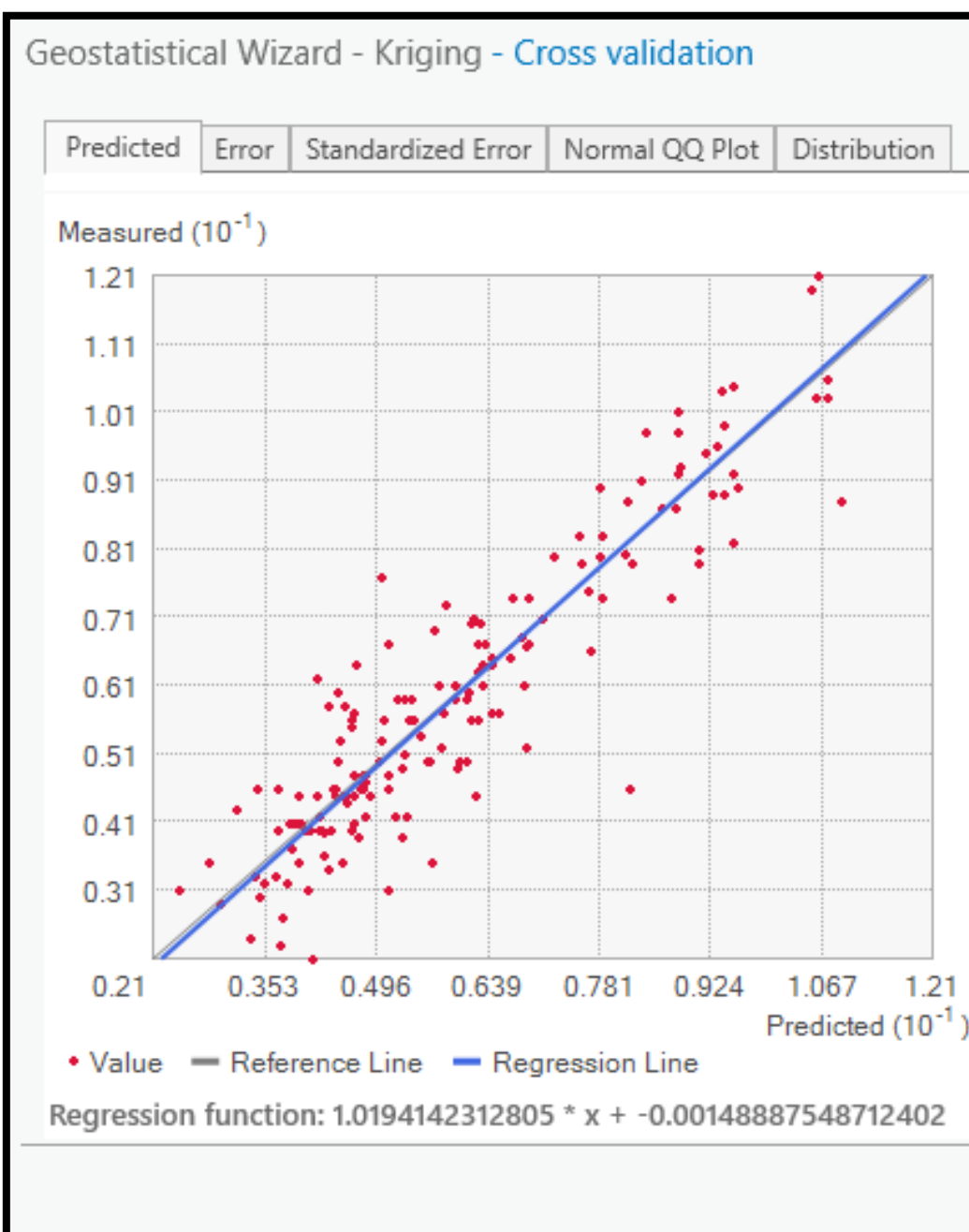


- Inverse distance weighted (IDW) interpolation explicitly makes the assumption that things that are close to one another are more alike than those that are farther apart.
- To predict a value for any unmeasured location, IDW uses the measured values surrounding the prediction location.
- IDW assumes that each measured point has a local influence that diminishes with distance. It gives greater weights to points closest to the prediction location, and the weights diminish as a function of distance, hence the name inverse distance weighted.
- Weights are proportional to the inverse of the distance (between the data point and the prediction location) raised to the power value p . As a result, as the distance increases, the weights decrease rapidly.

Result of IDW



Comparing Cross-Validation Results



Ordinary Kriging

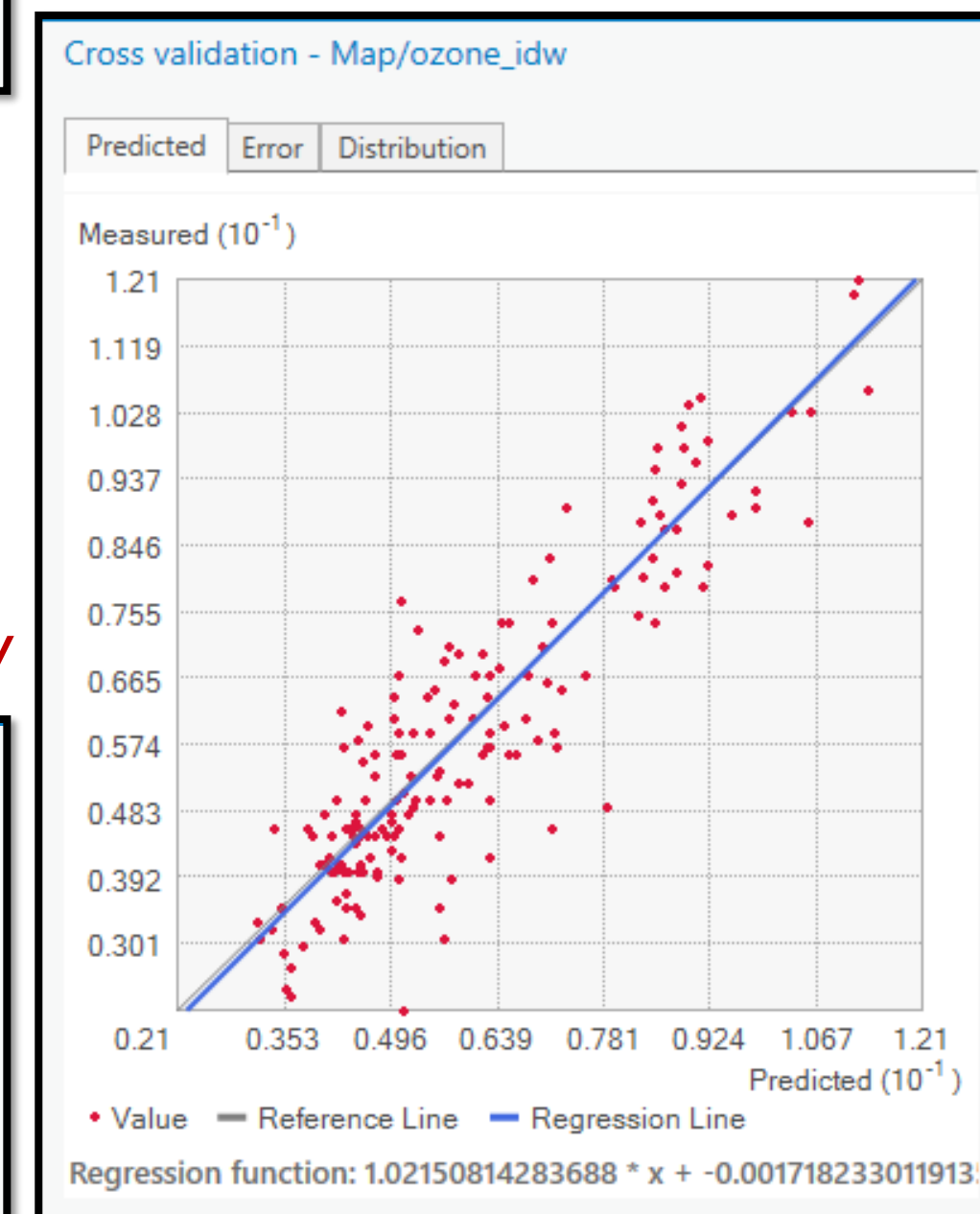
Summary Table

Count	167
Mean	0.000396454268027537
Root-Mean-Square	0.00862302755801119
Mean Standardized	0.0167857622534953
Root-Mean-Square Standardized	1.06511274997349
Average Standard Error	0.00860248117409058

IDW

Summary Table Method Report

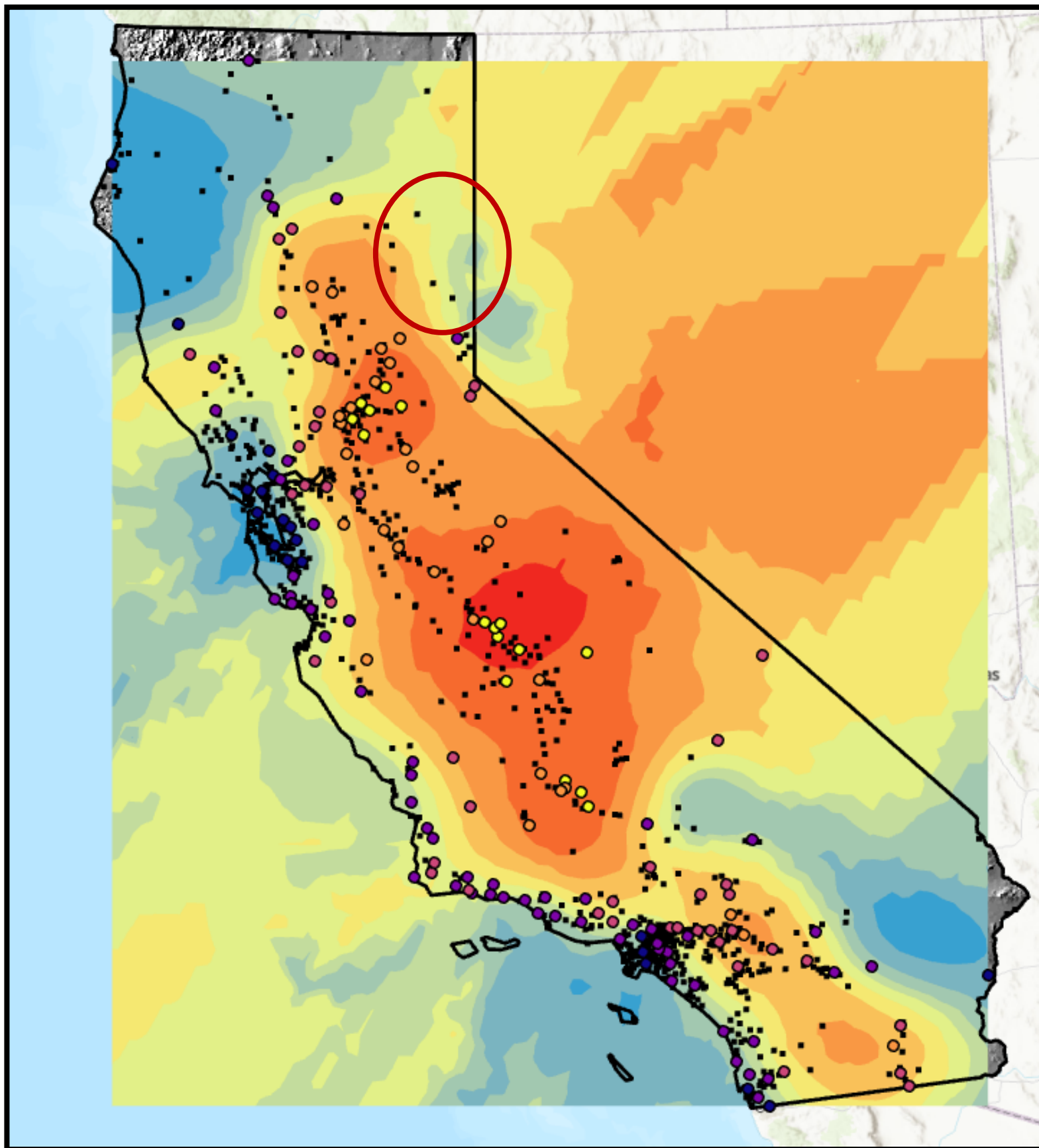
Count	167
Mean	0.000943651150445797
Root-Mean-Square	0.00944019656139528



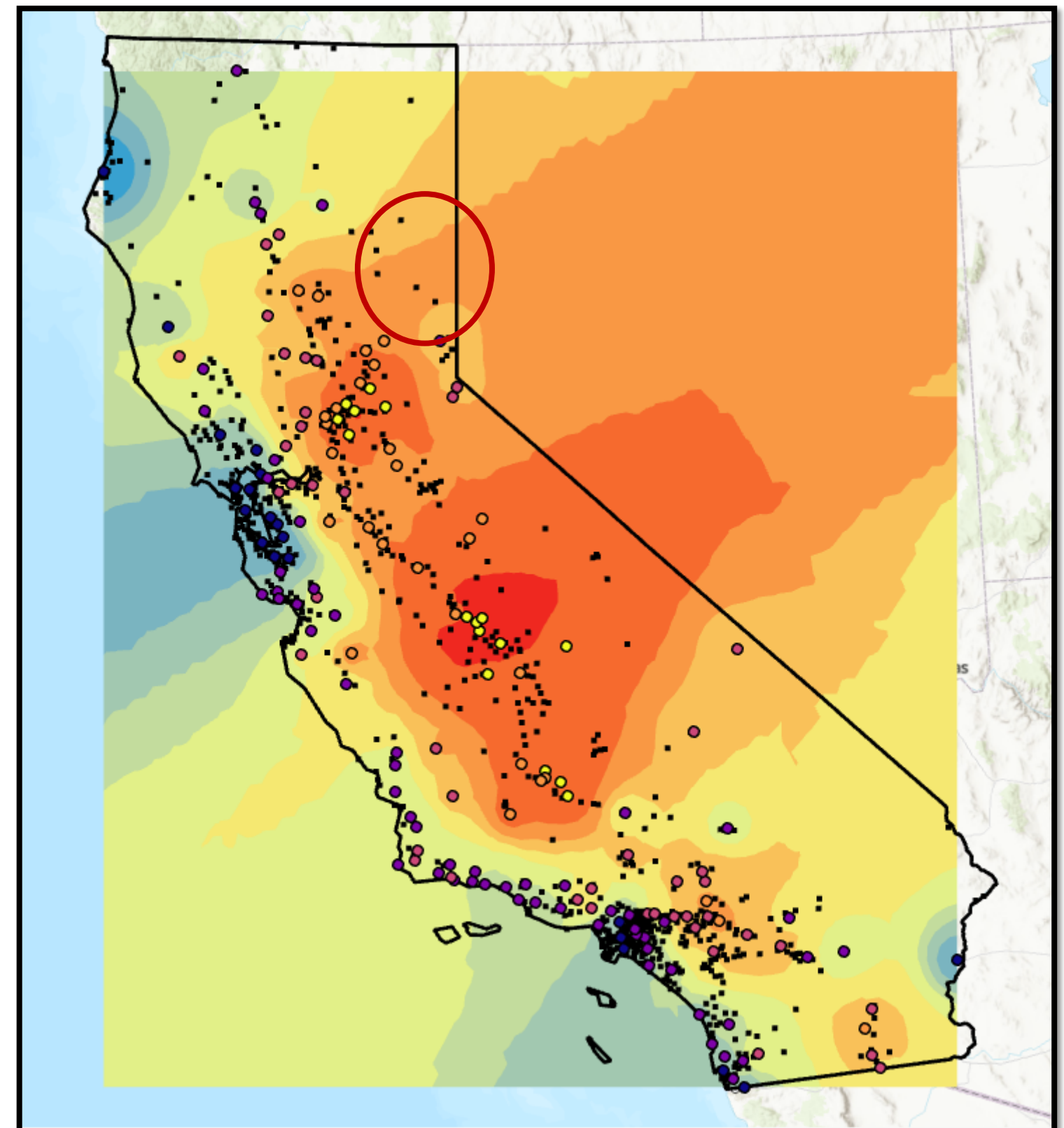
Comparing Results

Kriging captured local variations better than IDW.

Inverse Distance Weighted



Ordinary Kriging



Using Geostatistical Layer

Find the ozone concentration of cities by using geostatistical surfaces created by Kriging and IDW

Geoprocessing

GA Layer To Points

Parameters Environments

Input geostatistical layer
Kriging

Input point observation locations
ca_cities

Field to validate on

Output statistics at point locations
cities_kriging

☒ Append all fields from input features

Run

Catalog Geoprocessing Chart Properties

OBJECTID *	Shape *	AREANAME	Included	Predicted	Standard Error
1	Point	Acton	Yes	0.059412	0.003983
2	Point	Adelanto	Yes	0.059742	0.003842
3	Point	Agoura Hills	Yes	0.04514	0.003409
4	Point	Alameda	Yes	0.031042	0.002829
5	Point	Alamo	Yes	0.043938	0.002591
6	Point	Albany	Yes	0.032995	0.002666
7	Point	Alhambra	Yes	0.043182	0.002179
8	Point	Aliso Viejo	Yes	0.04107	0.003772
9	Point	Alondra Park	Yes	0.035284	0.003885
10	Point	Alpine	Yes	0.052331	0.004392

OBJECTID *	Shape *	AREANAME	Included	Predicted
1	Point	Acton	Yes	0.05425
2	Point	Adelanto	Yes	0.059709
3	Point	Agoura Hills	Yes	0.048418
4	Point	Alameda	Yes	0.032905
5	Point	Alamo	Yes	0.049158
6	Point	Albany	Yes	0.03373
7	Point	Alhambra	Yes	0.040969
8	Point	Aliso Viejo	Yes	0.044995
9	Point	Alondra Park	Yes	0.037758
10	Point	Alpine	Yes	0.056956

Geoprocessing

GA Layer To Points

Parameters Environments

Input geostatistical layer
ozone_idw

Input point observation locations
ca_cities

Field to validate on

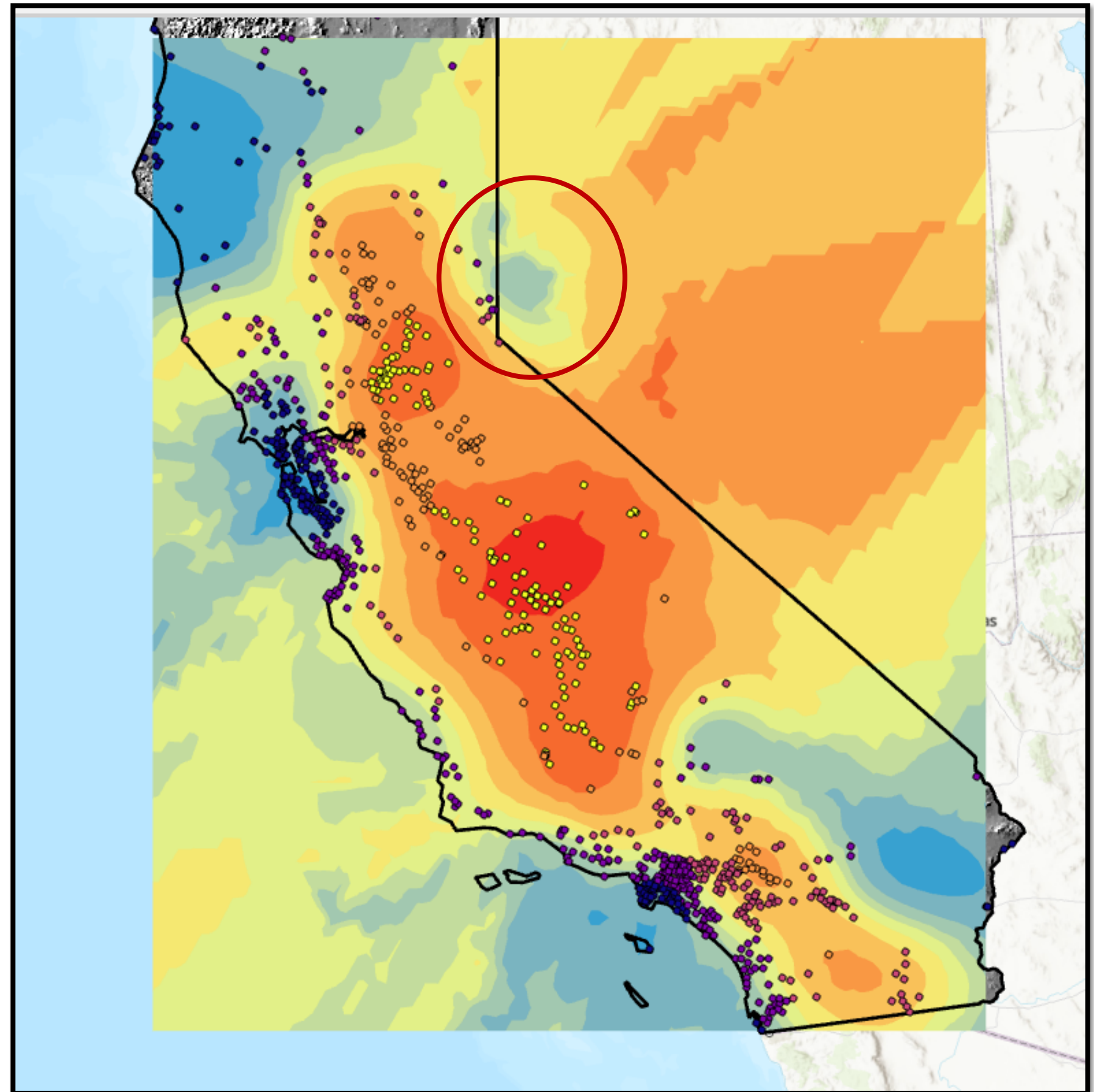
Output statistics at point locations
cities_idw

☒ Append all fields from input features

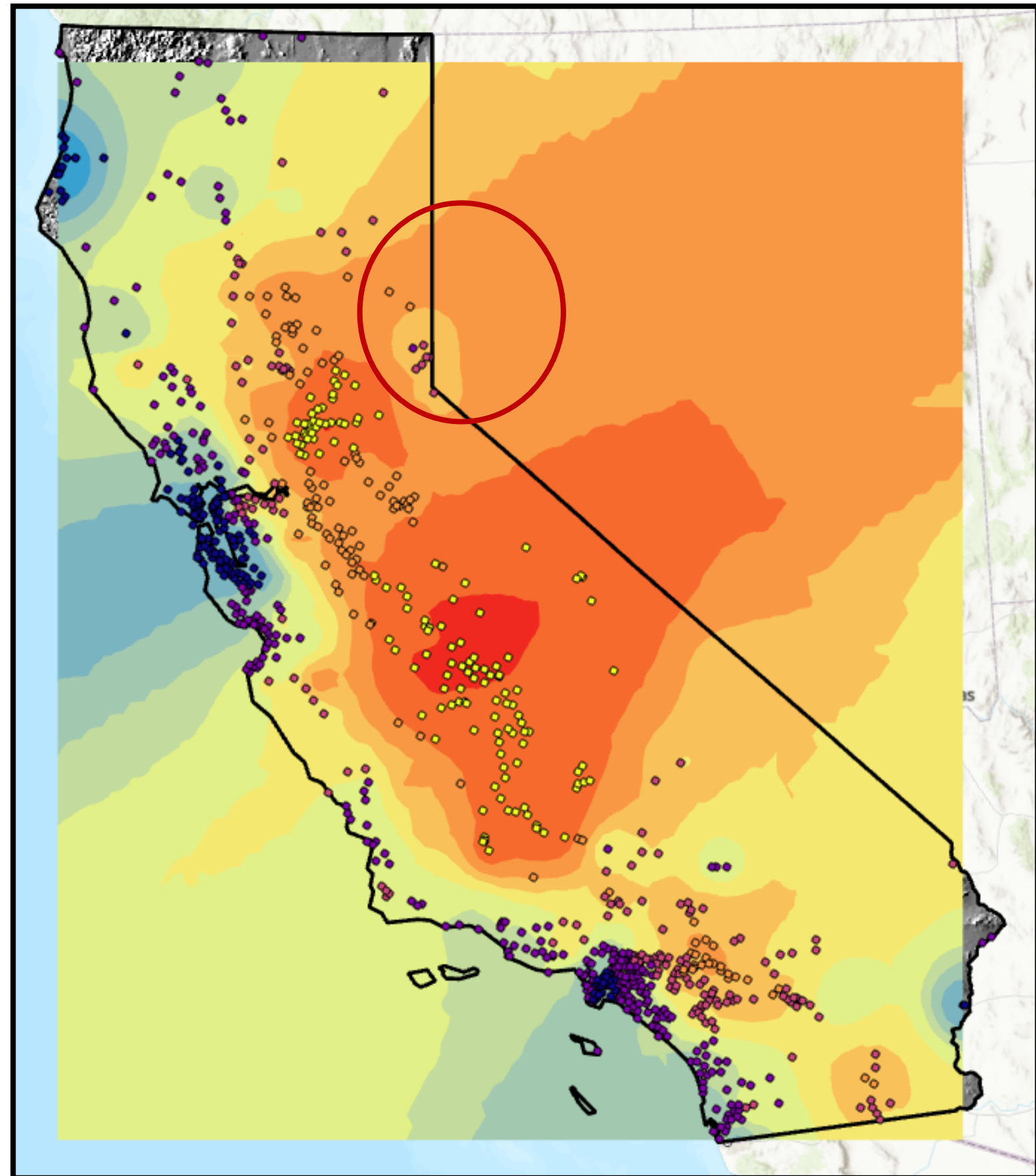
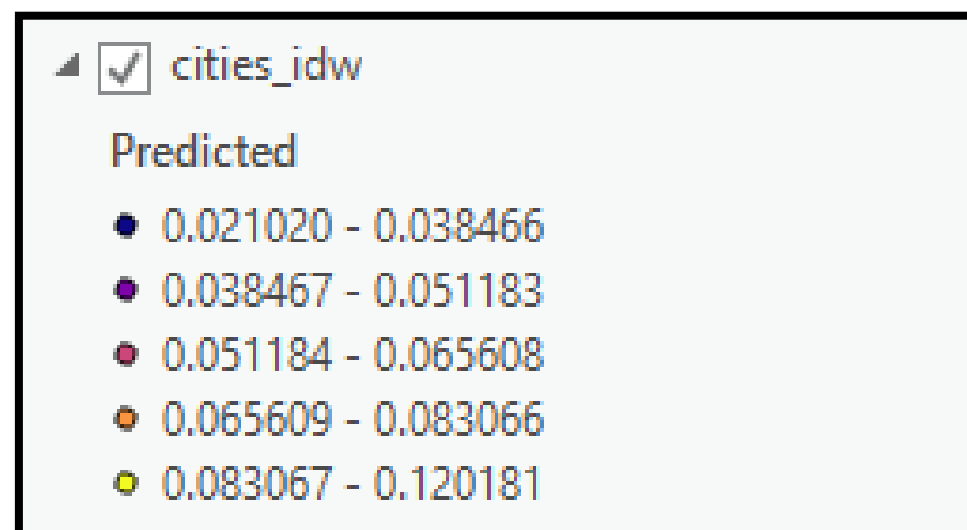
Run

Catalog Geoprocessing Chart Properties

Ozone Concentration of Cities Ordinary Kriging



Ozone Concentration of Cities IDW



Comparing Values

Compare the results for the area represented as circle in previous slides

cities_kriging					cities_idw					
Field:	Selection:	Highlighted:			Field:	Selection:	Highlighted:			
OBJECTID *	Shape *	AREANAME	Included	Predicted	OBJECTID *	Shape *	AREANAME	Included	Predicted	
209	Point	Dollar Point	Yes	0.054139	209	Point	Dollar Point	Yes	0.060467	
300	Point	Glenshire-Devonshire	Yes	0.047254	300	Point	Glenshire-Devonshire	Yes	0.052592	
378	Point	Kings Beach	Yes	0.048796	378	Point	Kings Beach	Yes	0.059347	
458	Point	Loyalton	Yes	0.04727	458	Point	Loyalton	Yes	0.069461	
616	Point	Portola	Yes	0.053769	616	Point	Portola	Yes	0.072397	
745	Point	South Lake Tahoe	Yes	0.059933	745	Point	South Lake Tahoe	Yes	0.058046	
764	Point	Sunnyside-Tahoe City	Yes	0.058212	764	Point	Sunnyside-Tahoe City	Yes	0.063709	
772	Point	Tahoe Vista	Yes	0.049889	772	Point	Tahoe Vista	Yes	0.057692	
792	Point	Truckee	Yes	0.052265	792	Point	Truckee	Yes	0.049161	
Click to add new row.					Click to add new row.					

Aim of the Study:

- *Ozone concentration is measured at monitoring stations throughout the state of California. The locations of monitoring stations and concentration levels of ozone are known for all the stations, but the ozone values for other (unmonitored) locations in California are also of interest.*

Output Data:

- *Results of Kriging and IDW (Raster)*
- *Cities having Ozone Concentration for Kriging and IDW (Vector-Point)*
- *Cross-validation results of Kriging and IDW (Vector-Point)*



Contact:

akinom@itu.edu.tr