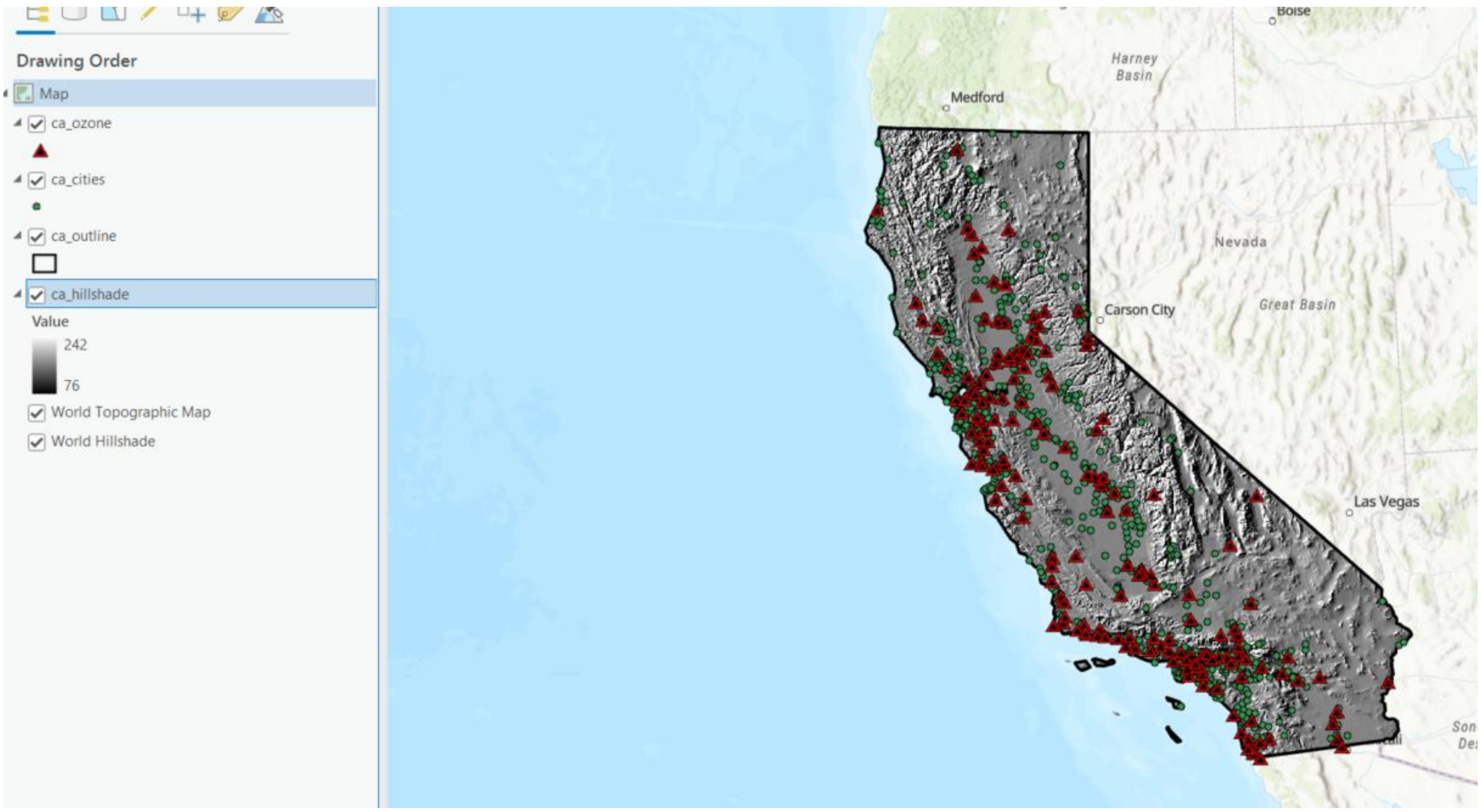
Ozone Concentration Mapping Using Geostatistical Methods in California

Aim of the Study

• To estimate ozone concentration levels at unmonitored locations across California.
• To create continuous prediction surfaces using geostatistical interpolation methods.
• To compare the performance of Kriging and Inverse Distance Weighted (IDW) techniques.

Input Data

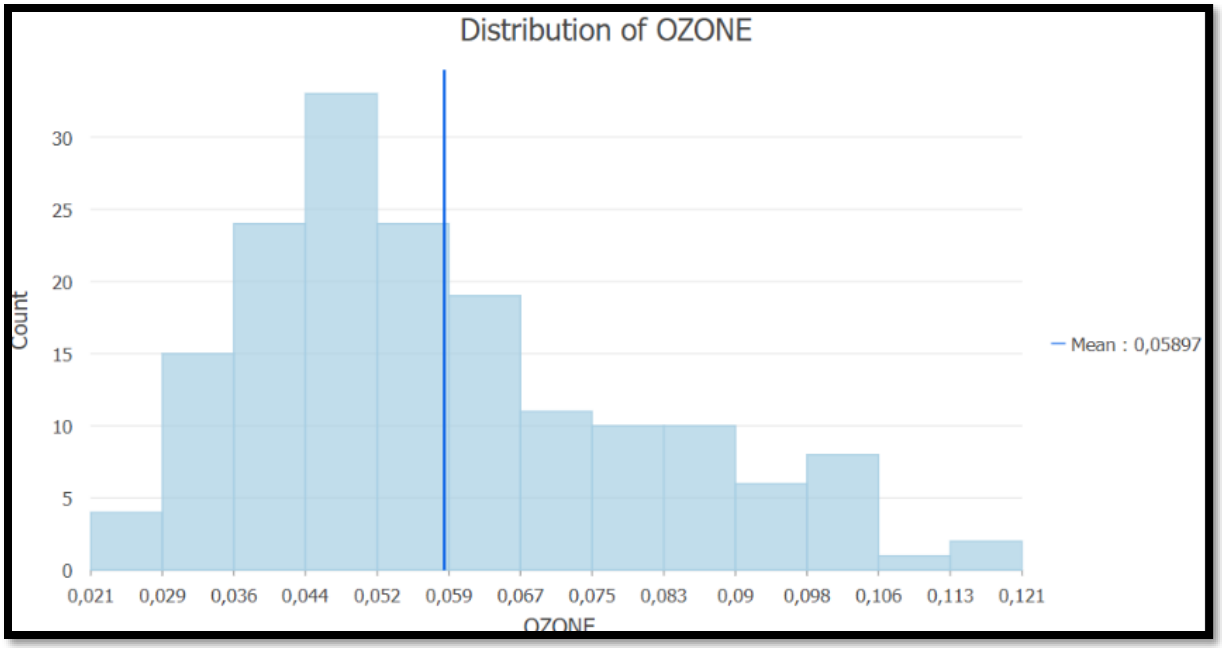
• Ozone Concentration (measured on September 6, 2007, between 3:00 and 4:00 p.m.) - Vector (Point)
• City Centers - Vector (Point)
• California State Border - Vector (Polygon)
• Hillshade - Raster

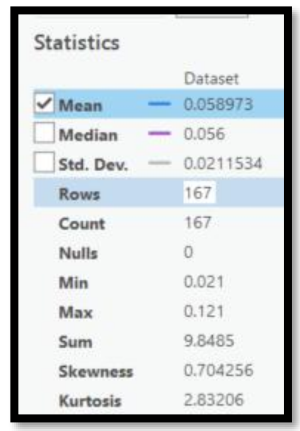


Methodology Overview

1. Exploratory Data Analysis (Statistics and Histogram)
2. Geostatistical Analyst Tools
   - Ordinary Kriging
   - Inverse Distance Weighted (IDW)
3. Semivariogram Modeling (for Kriging)
4. Prediction Surface Generation
5. Cross-Validation for Model Performance Assessment

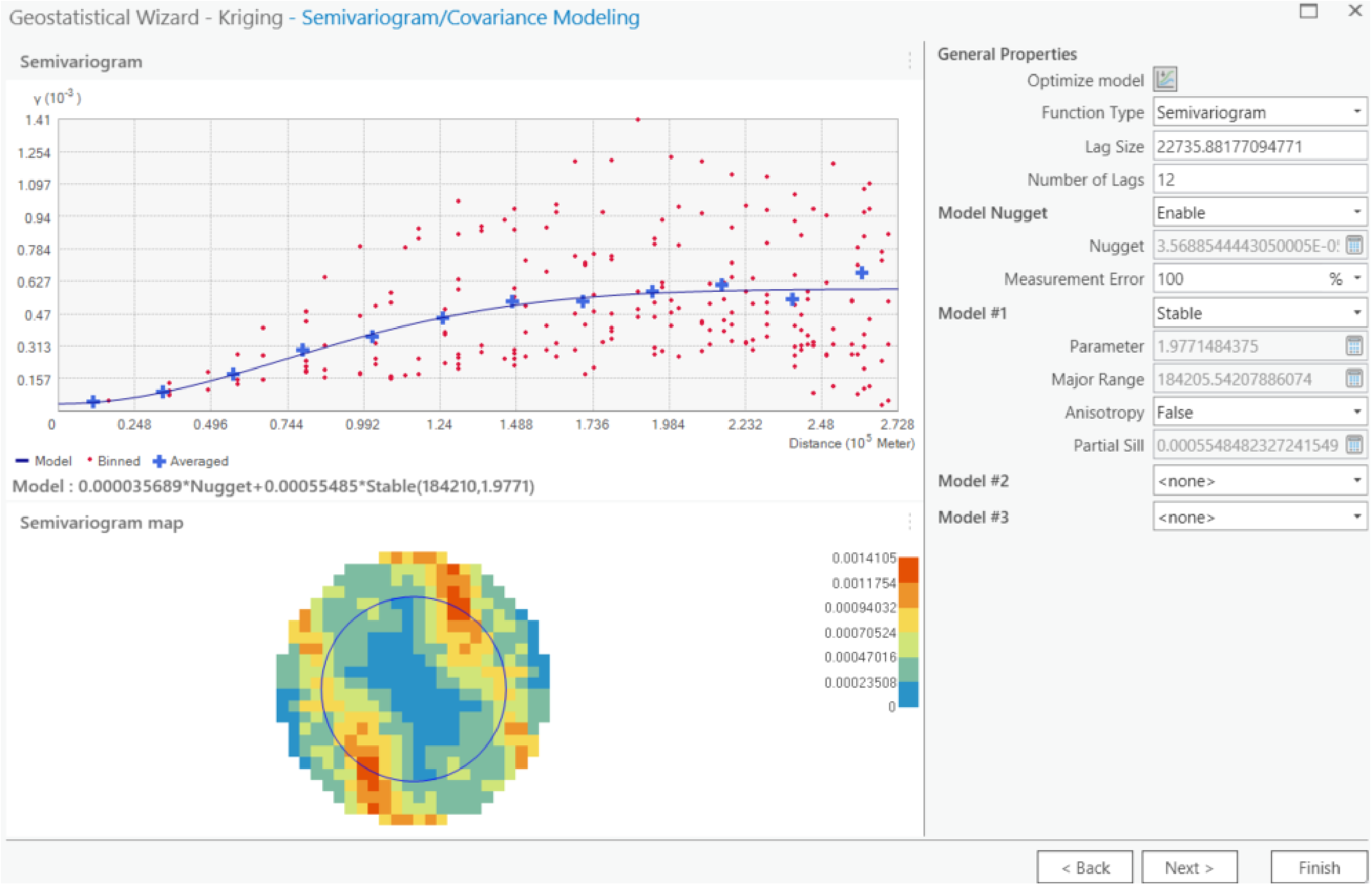
[Insert infographic comparing Kriging and IDW methods]





As an initial check, if the **mean and median values are close**, it suggests the data might follow a **normal distribution**.

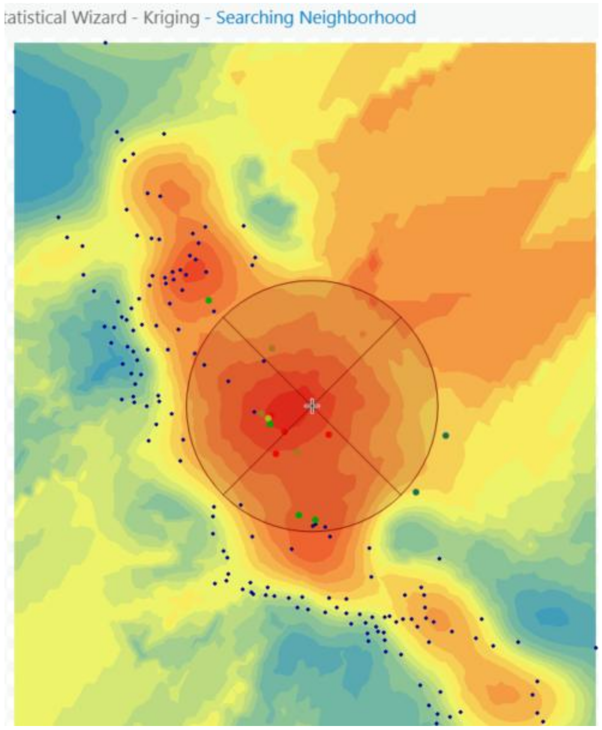
However, the **ozone data histogram** shows that the distribution is **unimodal** (a single peak) and **right-skewed**. The extended right tail indicates the presence of **a few sample points with high ozone concentration values**. This suggests that the data **does not closely follow a normal distribution**.



Spatial correlation in ozone concentrations exists up to **184 km**. **Low nugget** indicates minimal measurement error. The **Stable semi variogram model** fits the data well, capturing the underlying spatial structure. No anisotropy was considered in this model, assuming uniform spatial behavior across directions.

Kriging Results

• A semivariogram model was fitted to capture the spatial relationships between the measured points.
• The prediction surface was generated using Ordinary Kriging.
• Cross-validation was performed to evaluate prediction accuracy.



- The **predicted value at an unmeasured location** is generally **most similar to the values of nearby measured points**.
- In the diagram, the **red points** contribute **more weight** to the prediction than the **green points**, because they are **closer** to the location being estimated.
- By applying the **semivariogram/covariance model** developed earlier, along with the surrounding data points, it is possible to **predict the value at the unknown location**.

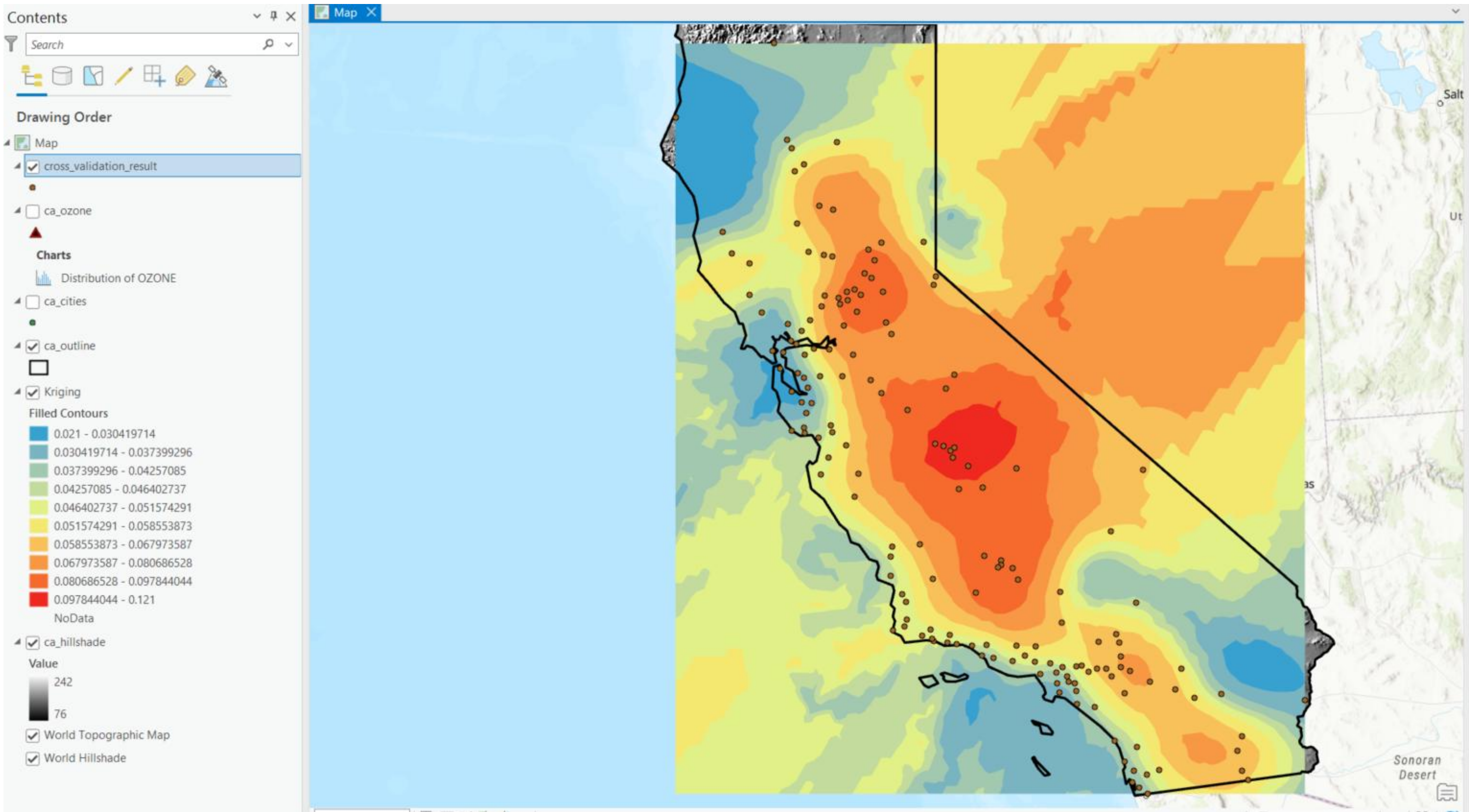
Summary | Table

Count	167
Mean	0.000396454268027538
Root-Mean-Square	0.00862302755801119
Mean Standardized	0.0167857622534953
Root-Mean-Square Standardized	1.06511274997349
Average Standard Error	0.00860248117409058

The **cross-validation results** indicate that the **Kriging model** performs well:

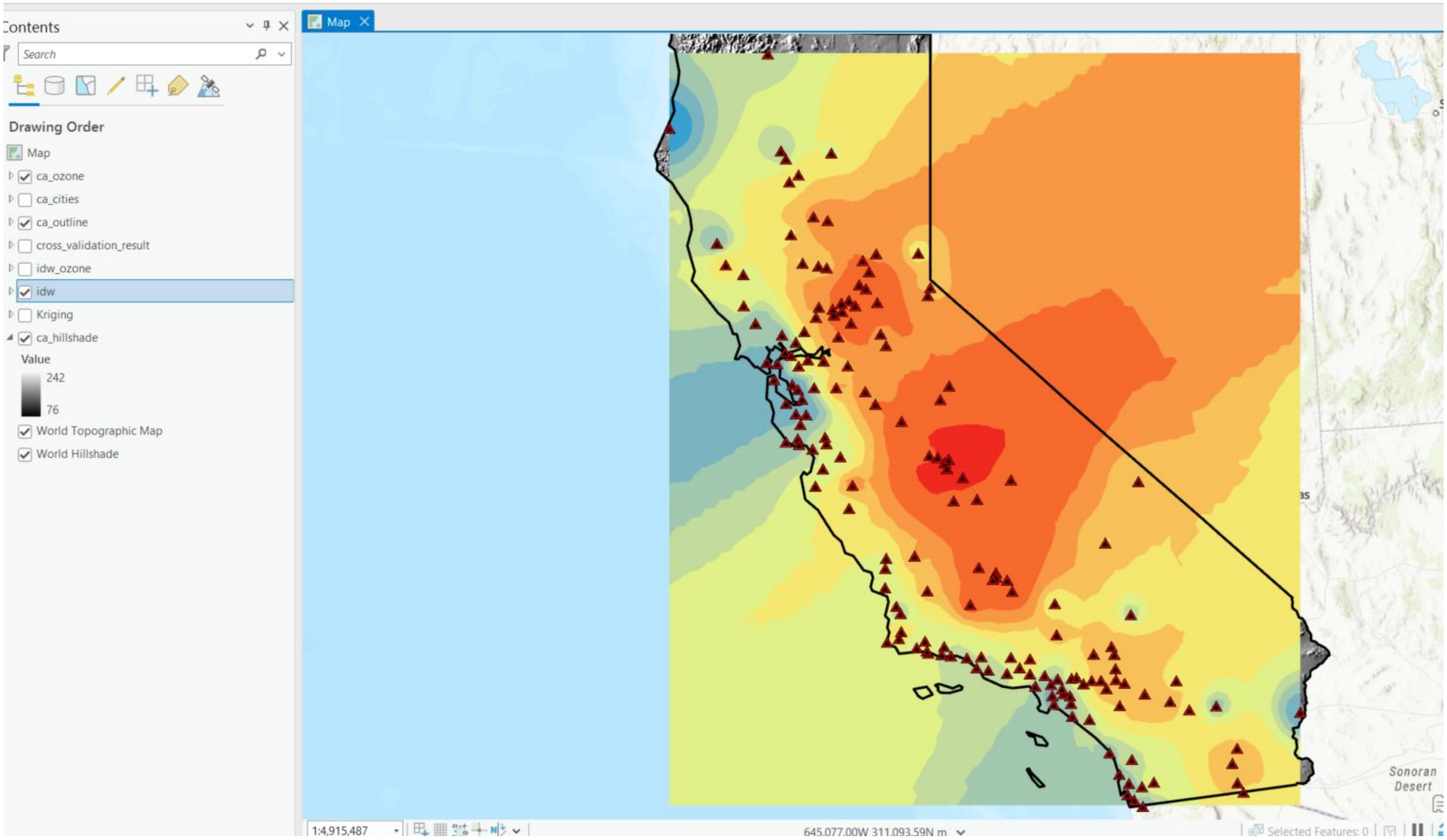
- The **Mean error** is very close to **zero**, suggesting **no bias** in predictions.
- The **Root Mean Square Error (RMSE)** reflects the average prediction error, and its alignment with the **Average Standard Error** shows that the model estimates its uncertainty accurately.
- The **Standardized RMSE (1.065)** being close to **1** indicates that the prediction errors are consistent with the estimated standard errors.

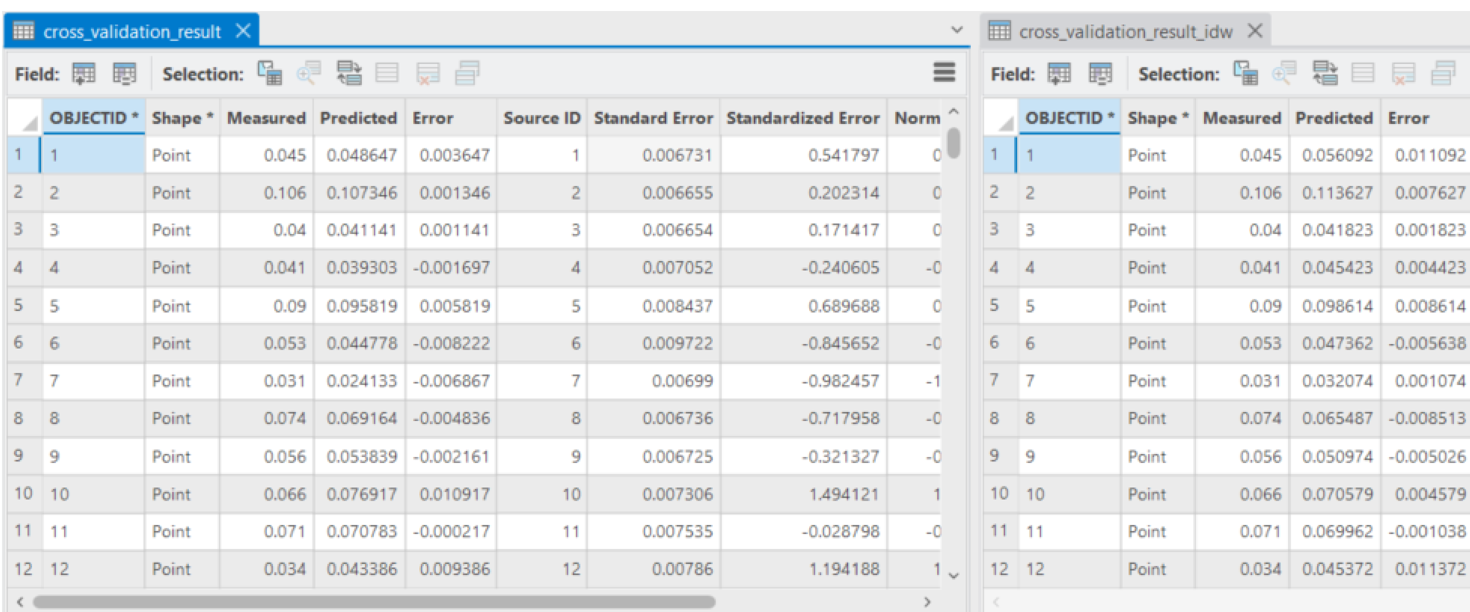
Overall, these metrics suggest that the model provides **reliable and unbiased predictions**.


Contents
Search
Drawing Order
Map
cross_validation_result
ca_ozone
Charts
Distribution of OZONE
ca_cities
ca_outline
Kriging
Filled Contours
0.021 - 0.030419714
0.030419714 - 0.037399296
0.037399296 - 0.04257085
0.04257085 - 0.046402737
0.046402737 - 0.051574291
0.051574291 - 0.058553873
0.058553873 - 0.067973587
0.067973587 - 0.080686528
0.080686528 - 0.097844044
0.097844044 - 0.121
NoData
ca_hillshade
Value
242
76
World Topographic Map
World Hillshade
Map
Salt
Sonoran
Desert

IDW Results

- IDW interpolation was applied, assuming that closer points have a stronger influence.
- A prediction surface was generated for ozone concentration.
- Cross-validation results were obtained for comparison.





cross_validation_result

	OBJECTID *	Shape *	Measured	Predicted	Error	Source ID	Standard Error	Standardized Error	Norm
1	1	Point	0.045	0.048647	0.003647	1	0.006731	0.541797	0
2	2	Point	0.106	0.107346	0.001346	2	0.006655	0.202314	0
3	3	Point	0.04	0.041141	0.001141	3	0.006654	0.171417	0
4	4	Point	0.041	0.039303	-0.001697	4	0.007052	-0.240605	-0
5	5	Point	0.09	0.095819	0.005819	5	0.008437	0.689688	0
6	6	Point	0.053	0.044778	-0.008222	6	0.009722	-0.845652	-0
7	7	Point	0.031	0.024133	-0.006867	7	0.00699	-0.982457	-1
8	8	Point	0.074	0.069164	-0.004836	8	0.006736	-0.717958	-0
9	9	Point	0.056	0.053839	-0.002161	9	0.006725	-0.321327	-0
10	10	Point	0.066	0.076917	0.010917	10	0.007306	1.494121	1
11	11	Point	0.071	0.070783	-0.000217	11	0.007535	-0.028798	-0
12	12	Point	0.034	0.043386	0.009386	12	0.00786	1.194188	1

cross_validation_result_idw

	OBJECTID *	Shape *	Measured	Predicted	Error
1	1	Point	0.045	0.056092	0.011092
2	2	Point	0.106	0.113627	0.007627
3	3	Point	0.04	0.041823	0.001823
4	4	Point	0.041	0.045423	0.004423
5	5	Point	0.09	0.098614	0.008614
6	6	Point	0.053	0.047362	-0.005638
7	7	Point	0.031	0.032074	0.001074
8	8	Point	0.074	0.065487	-0.008513
9	9	Point	0.056	0.050974	-0.005026
10	10	Point	0.066	0.070579	0.004579
11	11	Point	0.071	0.069962	-0.001038
12	12	Point	0.034	0.045372	0.011372

Kriging is the more reliable interpolation method for this dataset, offering both **higher accuracy** and **uncertainty quantification**, which are crucial for environmental analyses like ozone concentration mapping.

Comparison of Methods

- The performance of Kriging and IDW methods were compared using cross-validation results.
- Kriging demonstrated better accuracy in capturing local variations compared to IDW.

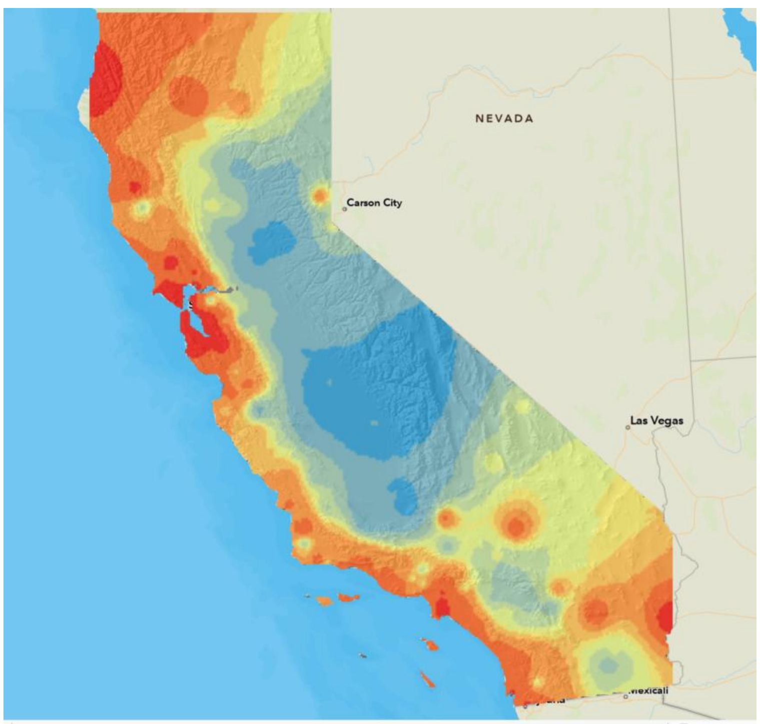
According to city location determined ozone values compared.

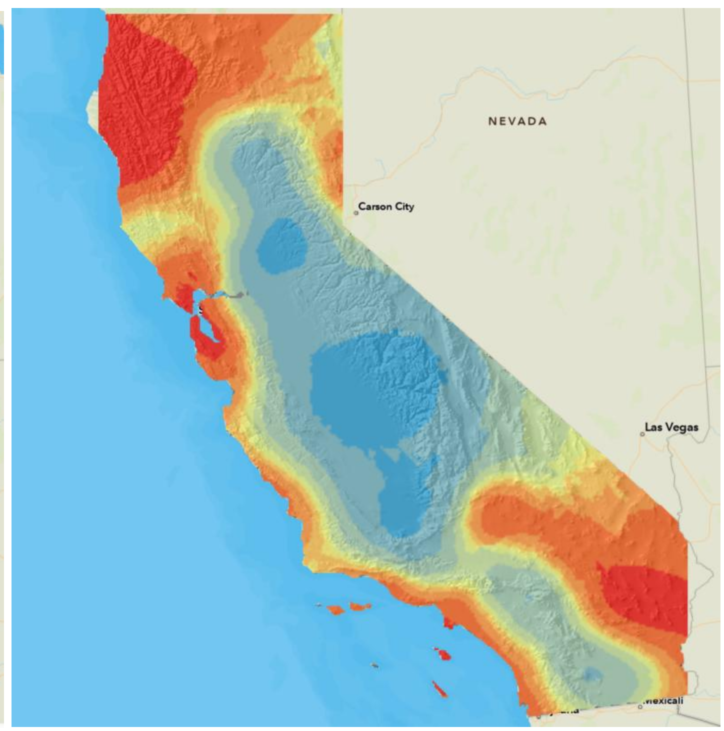
GALayerToPoints_idw

	OBJECTID *	Shape *	AREANAME	Included	Predicted
1	1	Point	Acton	Yes	0.05425
2	2	Point	Adelanto	Yes	0.059709
3	3	Point	Agoura Hills	Yes	0.048418
4	4	Point	Alameda	Yes	0.032905
5	5	Point	Alamo	Yes	0.049158
6	6	Point	Albany	Yes	0.03373
7	7	Point	Alhambra	Yes	0.040969
8	8	Point	Aliso Viejo	Yes	0.044995
9	9	Point	Alondra Park	Yes	0.037758
10	10	Point	Alpine	Yes	0.056956
11	11	Point	Altadena	Yes	0.042586
12	12	Point	Alta Sierra	Yes	0.08583
13	13	Point	Alturas	Yes	0.057343
14	14	Point	Amador City	Yes	0.080002

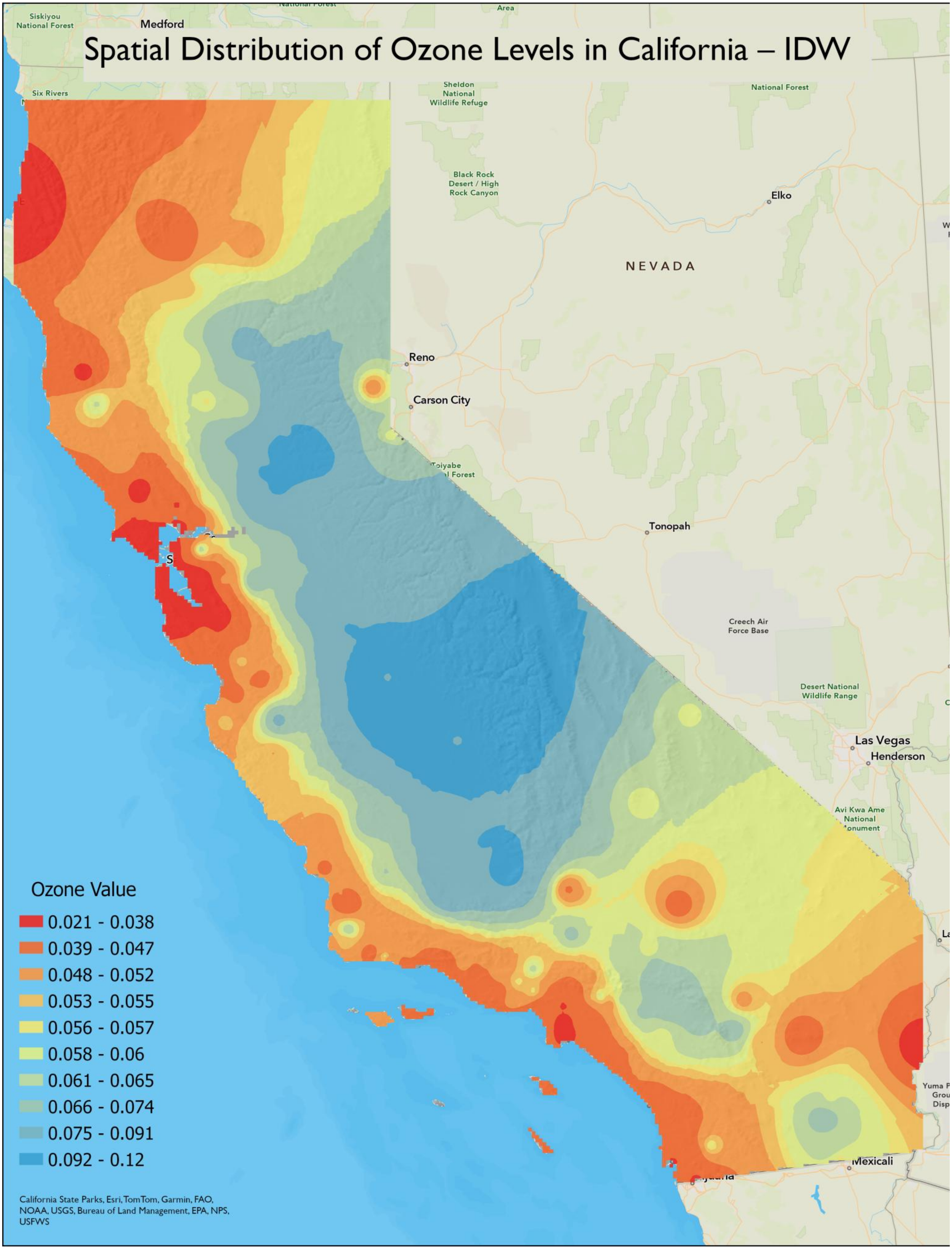
GALayerToPoints1_krigging

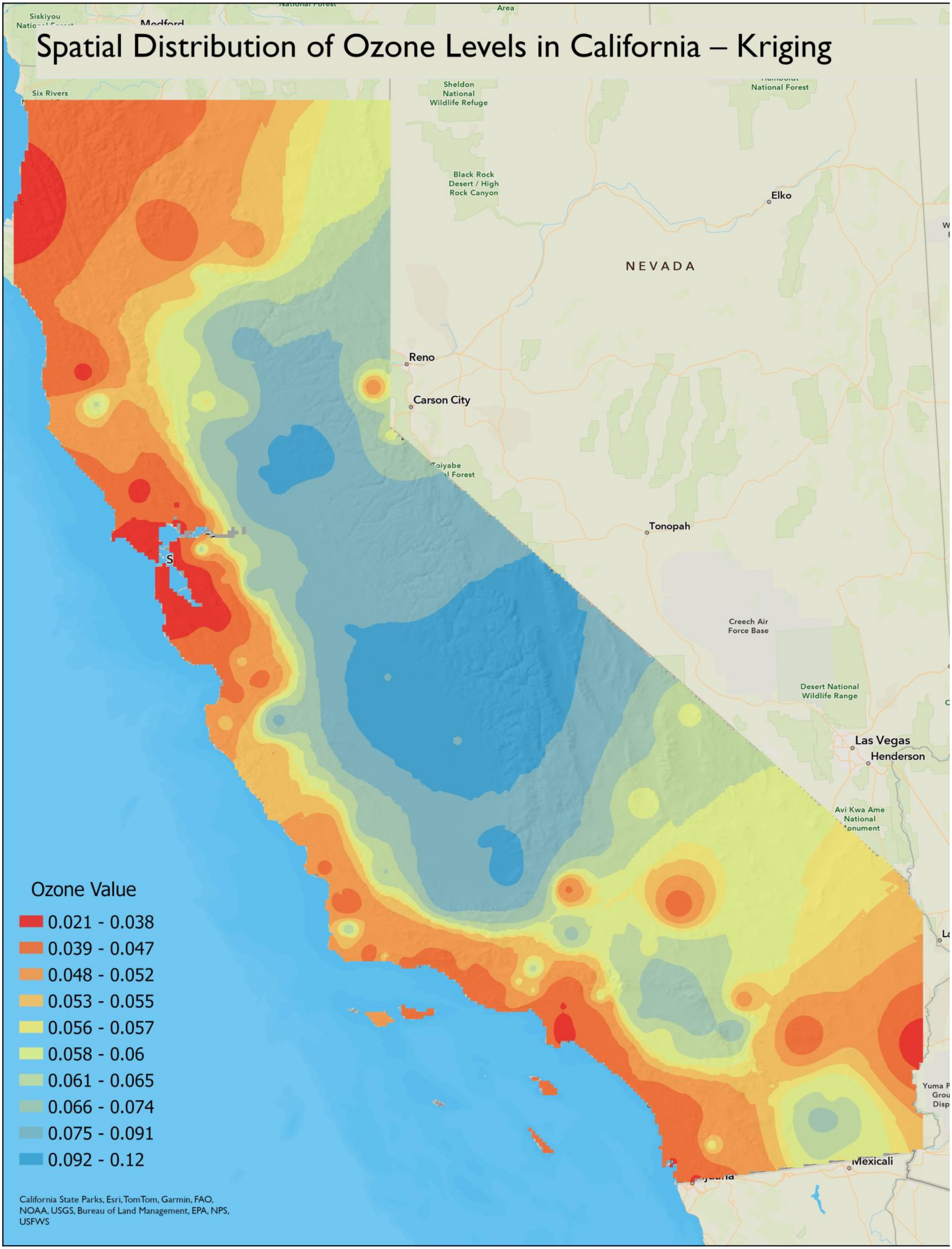
	OBJECTID *	Shape *	AREANAME	Included	Predicted	Standard Error
1	1	Point	Acton	Yes	0.059412	0.003983
2	2	Point	Adelanto	Yes	0.059742	0.003842
3	3	Point	Agoura Hills	Yes	0.04514	0.003409
4	4	Point	Alameda	Yes	0.031042	0.002829
5	5	Point	Alamo	Yes	0.043938	0.002591
6	6	Point	Albany	Yes	0.032995	0.002666
7	7	Point	Alhambra	Yes	0.043182	0.002179
8	8	Point	Aliso Viejo	Yes	0.04107	0.003772
9	9	Point	Alondra Park	Yes	0.035284	0.003885
10	10	Point	Alpine	Yes	0.052331	0.004392
11	11	Point	Altadena	Yes	0.04861	0.002754
12	12	Point	Alta Sierra	Yes	0.085084	0.002882
13	13	Point	Alturas	Yes	0.046655	0.027079
14	14	Point	Amador City	Yes	0.086605	0.003513





IDW Kriging


Spatial Distribution of Ozone Levels in California – IDW
Siskiyou National Forest
Medford
Six Rivers
Sheldon National Wildlife Refuge
National Forest
Black Rock Desert / High Rock Canyon
Elko
NEVADA
Reno
Carson City
Toiyabe
Tonopah
Creech Air Force Base
Desert National Wildlife Range
Las Vegas
Henderson
Avi Kwa Ame National
Mexicali
Ozone Value
0.021 - 0.038
0.039 - 0.047
0.048 - 0.052
0.053 - 0.055
0.056 - 0.057
0.058 - 0.06
0.061 - 0.065
0.066 - 0.074
0.075 - 0.091
0.092 - 0.12
California State Parks, Esri, TomTom, Garmin, FAO, NOAA, USGS, Bureau of Land Management, EPA, NPS, USFWS


Spatial Distribution of Ozone Levels in California – Kriging
Ozone Value
0.021 - 0.038
0.039 - 0.047
0.048 - 0.052
0.053 - 0.055
0.056 - 0.057
0.058 - 0.06
0.061 - 0.065
0.066 - 0.074
0.075 - 0.091
0.092 - 0.12
NEVADA
Reno
Carson City
Tonopah
Elko
Las Vegas
Henderson
Mexicali
Sheldon National Wildlife Refuge
Black Rock Desert / High Rock Canyon
Creech Air Force Base
Desert National Wildlife Range
Avi Kwa Ame National Monument
California State Parks, Esri, TomTom, Garmin, FAO, NOAA, USGS, Bureau of Land Management, EPA, NPS, USFWS

Conclusions & Evaluation

In this study, two different spatial interpolation techniques were applied to predict ozone concentration levels at unmeasured locations: **Ordinary Kriging** and **Inverse Distance Weighted (IDW)**. While both methods aim to estimate values based on known sample points, they differ significantly in their underlying assumptions and modeling approaches.

Kriging is a **geostatistical method** that models the **spatial autocorrelation** among data points using a **semi variogram**. This model captures the degree of similarity between nearby points and adjusts the weights accordingly. Kriging not only provides predictions but also estimates the **uncertainty (standard error)** associated with these predictions, making it a more robust approach when spatial patterns and dependencies are present.

In contrast, **IDW** is a **deterministic method** that relies solely on the **distance between sample points and prediction locations**. It assumes that points closer to the prediction location have more influence than those farther away, without considering any spatial structure or trends in the data. IDW does not provide any measure of uncertainty, which can be a limitation for certain applications.

The **cross-validation results** indicate that **Kriging outperforms IDW** in terms of prediction accuracy for this dataset. Kriging's ability to incorporate spatial relationships allowed it to capture **local variations** more effectively, which is especially important for environmental phenomena like ozone concentration. In contrast, IDW showed higher error rates and struggled to represent local fluctuations.

Overall, **Kriging is recommended** for datasets with evident **spatial autocorrelation**, while **IDW** may be suitable for simpler applications where spatial structure is minimal or unknown.