

# ArcGIS for Geostatistical Analyst: An Introduction

Steve Lynch and Eric Krause Redlands, CA.

#### Outline

- What is geostatistics?
- What is Geostatistical Analyst?
- Spatial autocorrelation
- Geostatistical Wizard and geoprocessing tools
- Where is it used?
- Demonstrations
- Tips and Tricks
- Conclusion
- Questions

#### Sessions of note...

Tuesday

- ArcGIS for Geostatistical Analyst: An Introduction (Tues 8:30-9:45 SDCC Rm17B)
- Creating Surfaces from Various Data Sources (Tues 3:15-4:30 SDCC Rm09)
- Concepts and Applications of Kriging (Tues 3:15-4:30 SDCC Rm17B)
- Empirical Bayesian Kriging and EBK Regression Prediction Robust Kriging as GP Tools (Tues 5:30-6:15 Th07)

#### Wednesday

- Choosing the Best Kriging Model for Your Data (Wed 11:30-12:15 SDCC Th07)
- ArcGIS for Geostatistical Analyst: An Introduction (Wed 1:30-2:45 SDCC Rm17B)

#### Thursday

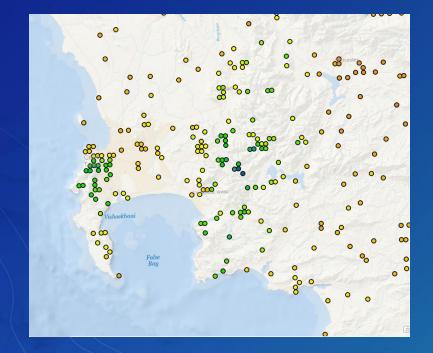
- Geostatistics in Practice: Learning Kriging Through Examples (Thurs 8:30-9:45 SDCC Rm10)
- Surface Interpolation in ArcGIS (Thurs 10:30-11:15 SDCC Th07)
- Performing Polygon-to-Polygon Predictions using Areal Interpolation (Thurs 11:30-12:15 Th07)
- Creating Surfaces from Various Data Sources (Thurs 3:15-4:30 SDCC Rm09)

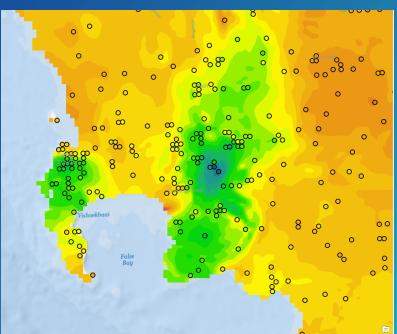
#### What is geostatistics?

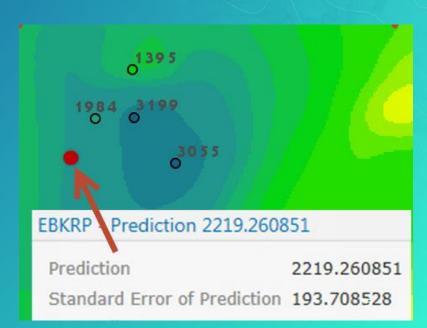
- is a class of statistics used to analyze and predict values associated with spatial phenomena.
- it incorporates the spatial coordinates of the data
- Has evolved to not only provide
  - interpolated values, but also
  - measures of uncertainty

#### ArcGIS for Geostatistical Analyst

Too expensive to measure everywhere, however, we want to know values everywhere.

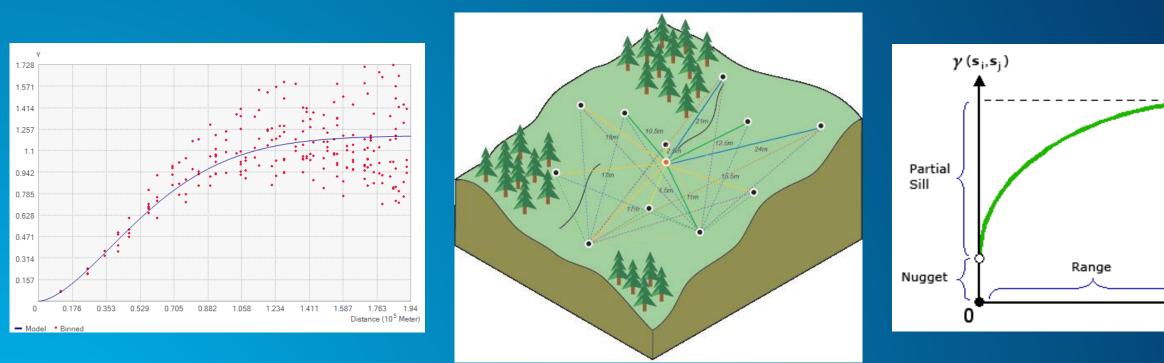






#### What is a semivariogram?

#### Semivariogram(distance h) = 0.5 \* average [ (value<sub>i</sub> – value<sub>i</sub>)<sup>2</sup>]

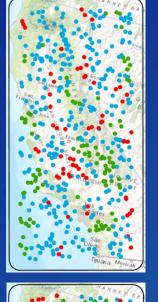


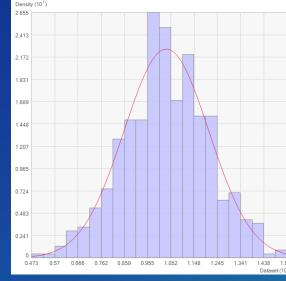
- Range = separation distance between pairs
- Sill = plateau the variogram reaches at the range

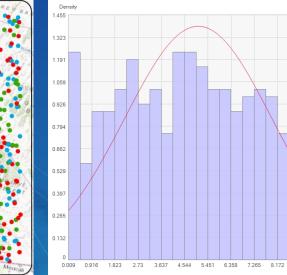
Sill

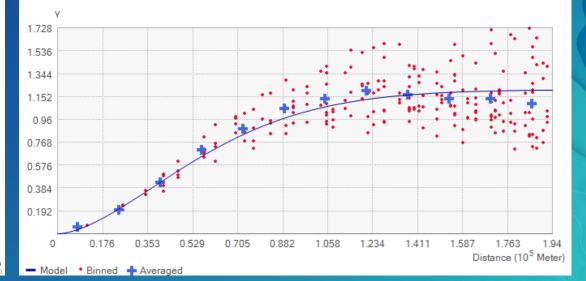
Nugget = sampling error and short scale variability

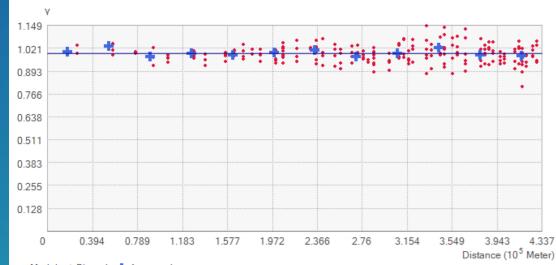
#### Spatial autocorrelation











9.079

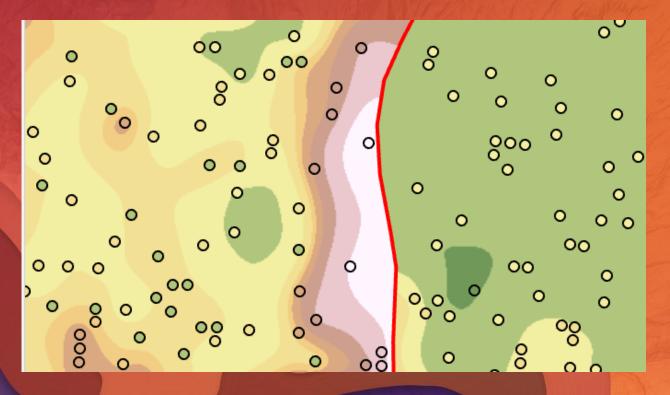
Dataset (10<sup>-1</sup>

9.986

## Geostatistical Analyst – What is it?

Provides a complete set of spatial analytical tools that range from techniques to explore the original data to postprocessing evaluation of data and predictions uncertainties.

- Geoprocessing tools
  - Use within ArcMap / Pro / Server
  - Modelbuilder
  - Scripting



# GP tool

Kernel Interpolation with Barriers Eric Krause

9

#### Geostatistical Analyst – Geoprocessing tools

Geostatistical Analyst Tools

#### 🔺 🚉 Interpolation

- Solution States Stat
- Section Prediction
- 🔨 Empirical Bayesian Kriging
- Slobal Polynomial Interpolation
- 🔨 IDW
- 🔨 Kernel Interpolation With Barriers
- 🔨 Local Polynomial Interpolation
- Noving Window Kriging
- Kadial Basis Functions

- 🔺 🔄 Sampling Network Design
  - Create Spatially Balanced Points
  - 🔨 Densify Sampling Network
- 🔺 🔄 Simulation
  - Stract Values To Table
  - 🔨 Gaussian Geostatistical Simulations
- 🔺 🚉 Utilities
  - 🔨 Cross Validation
  - Neighborhood Selection
  - 🔨 Semivariogram Sensitivity
  - 🔨 Subset Features

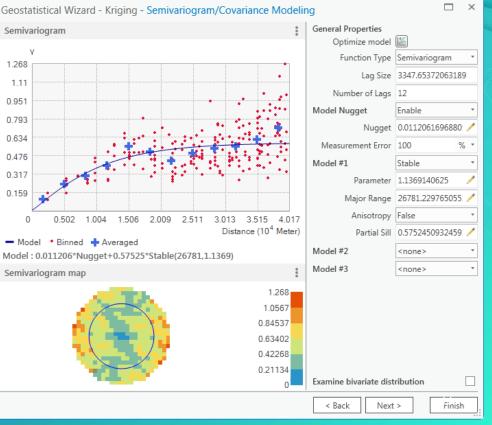
- Working with Geostatistical Layers
  - 🔨 Areal Interpolation Layer To Polygons
  - 🔨 Calculate Z-value
  - 🔨 Create Geostatistical Layer
  - 🔨 GA Layer To Contour
  - 🔨 GA Layer To Grid
  - 🔨 GA Layer To Points
  - GA Layer To Rasters
  - 🔨 Get Model Parameter
  - 🔨 Set Model Parameter

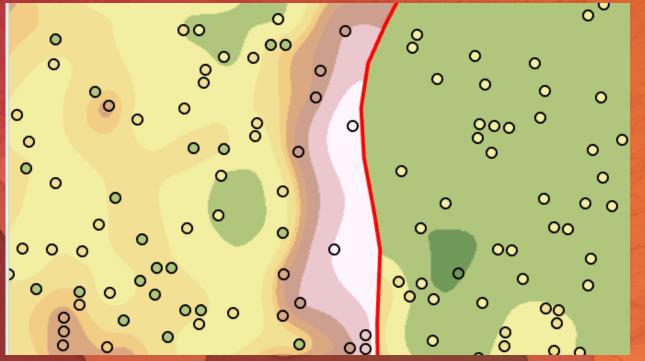
#### Geostatistical Analyst – What is it?

Provides a complete set of spatial analytical tools that range from techniques to explore the original data to postprocessing evaluation of data and predictions uncertainties.

Wizard

 is a dynamic set of pages that is designed to guide you through the process of constructing and evaluating the performance of an interpolation model.





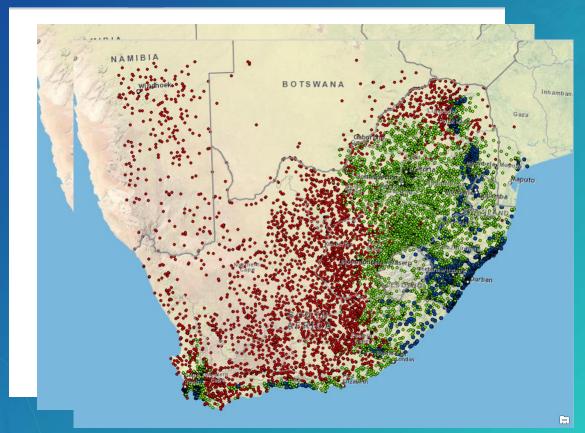
## **Geostatistical Wizard**

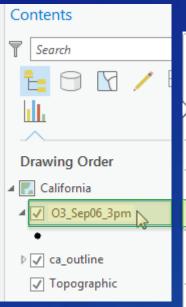
Kernel Interpolation with Barriers

Eric Krause

#### ESDA Exploratory Spatial Data Analysis

- Where is the data located?
- What are the values at the data points?
- How does the location of a point relate to its value?





F	Сору	
E×	Remove	
۲	Group	
2 IIII	Attribute Table	
	Design	Þ
	Joins and Relates	Þ
հ	Create Chart	Þ
Q.	Zoom To Layer	
Q	Zoom To Make Visible	
	Selection	Þ
R	Label	
æ	Labeling Properties	
1	Symbology	
<b></b>	Disable Pop-ups	
5	Configure Pop-ups	
	Data	Þ
<u> </u>	Save As Layer File	
÷	Share As Layer Package	
	Share As Web Layer	
	Overwrite Web Layer	
P	Properties	
	XXX	

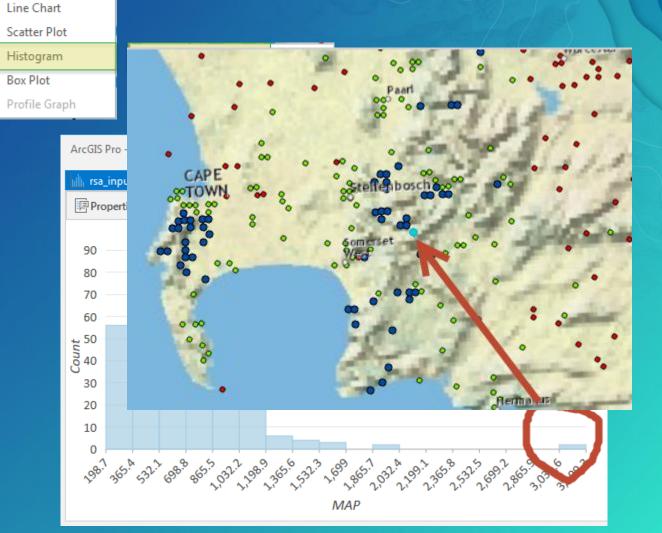
Bar Chart

dt.

đ

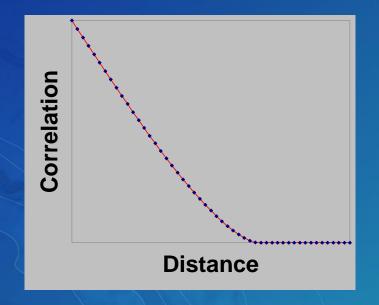
ŧŧî.

## Exploratory Spatial Data Analysis (ESDA)



## What is kriging?

- It is a geostatistical interpolation technique
- that models the spatial correlation of point measurements
- to estimate values at unmeasured locations.
- Associates uncertainty with the predictions



	Geostatistical Wizard - Kriging
	Ordinary Kriging
monto	<ul> <li>Prediction</li> <li>Quantile</li> <li>Probability</li> <li>Prediction Standard Error</li> </ul>
ements	Simple Kriging
	<ul> <li>Prediction</li> <li>Quantile</li> <li>Probability</li> <li>Prediction Standard Error</li> </ul>
	Universal Kriging
EBK Regression Prediction Empirical Bayesian Kriging	<ul> <li>Prediction</li> <li>Quantile</li> <li>Probability</li> <li>Prediction Standard Error</li> </ul>
	Indicator Kriging
	<ul> <li>Probability</li> <li>Standard Error of Indicators</li> </ul>
	Probability Kriging
	<ul> <li>Probability</li> <li>Standard Error of Indicators</li> </ul>
	Disjunctive Kriging
	<ul> <li>Prediction</li> <li>Probability</li> <li>Prediction Standard Error</li> <li>Standard Error of Indicators</li> </ul>

## What is kriging?

#### Tuesday

- Concepts and Applications of Kriging (Tues 3:15-4:30 SDCC Rm17B)
- EBK and EBK Regression Prediction Robust Kriging as GP Tools (Tues 5:30-6:15 Th07)

#### Wednesday

Choosing the Best Kriging Model for Your Data (Wed 11:30-12:15 SDCC Th07)

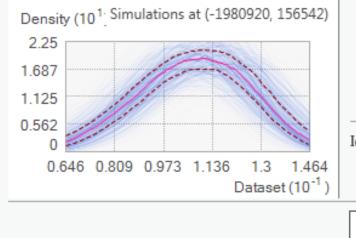
#### Thursday

- Geostatistics in Practice: Learning Kriging Through Examples (Thurs 8:30-9:45 SDCC Rm10)
- Performing Polygon-to-Polygon Predictions using Areal Interpolation (Thurs 11:30-12:15 Th07)

#### Geostatistical Wizard - Empirical Bayesian K

# Semivariograms Nuccet

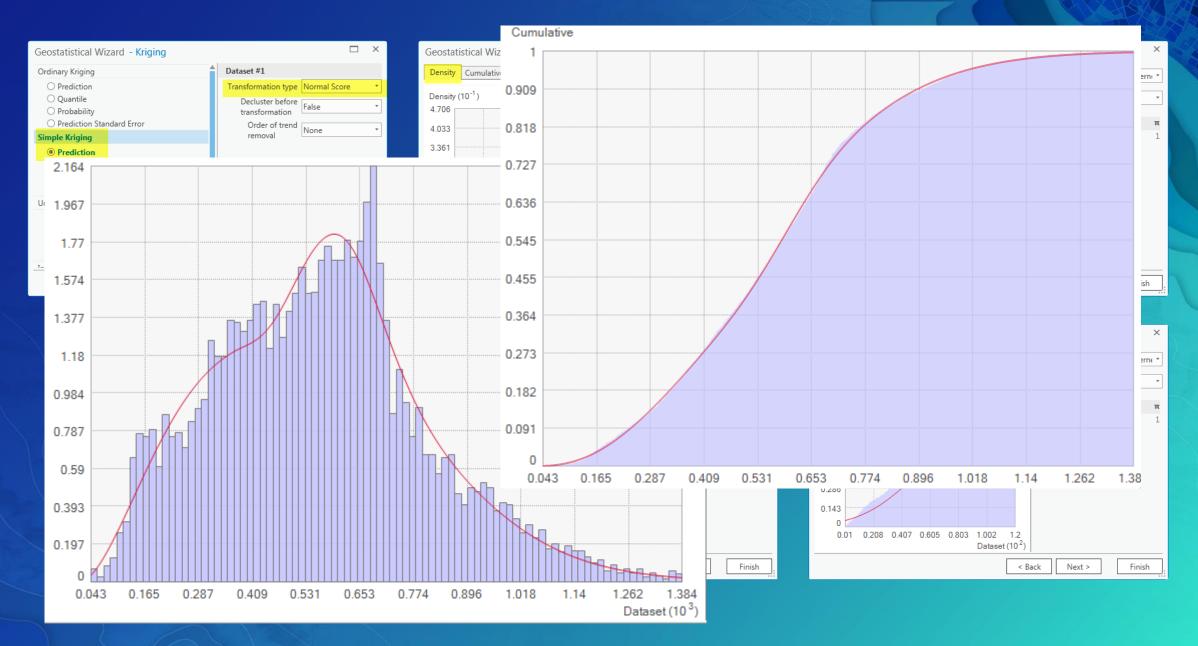
Seriivanogranis i ivugge	
Partial Sill Range Transforma	tion



Kriging		
ieneral Prope	rties	
Subse	et Size	100
Overlap	Factor	1
	per of ations	500
Output S Type	urface	Prediction •
Transform	nation	Empirical 🔹
Semivario Type	ogram	Whittle *
leighborhood	d Type	Standard Circular 🔻
Neig	imum Ihbors	15
	mum Ihbors	10
Secto	r Type	◯1 Sector ・
	Angle	0
F	Radius	90085.7106446951
dentify Resul	t	
	Х	-1980920
	v	156540
< Back	Nex	t > Finish

## **Geostatistical Wizard**

#### **Eric Krause**



More ESDA

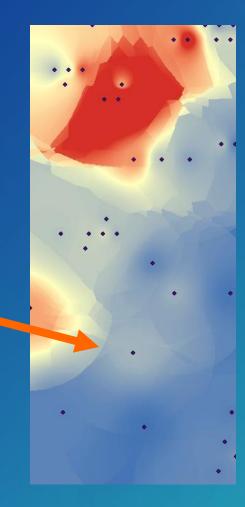
## Interpolation workflow

- ESDA
- Interpolate
- Goodness of fit

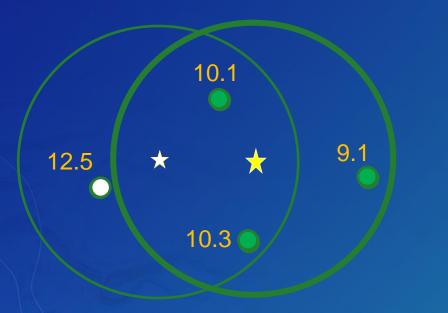
#### Why use ESRI's Geostatistical Analyst?

- Search neighborhood
  - Sectors
  - Smooth
- Chordal distance
- Cross validation
- Error maps
- Interactive Variography
- Barriers
- Simulations

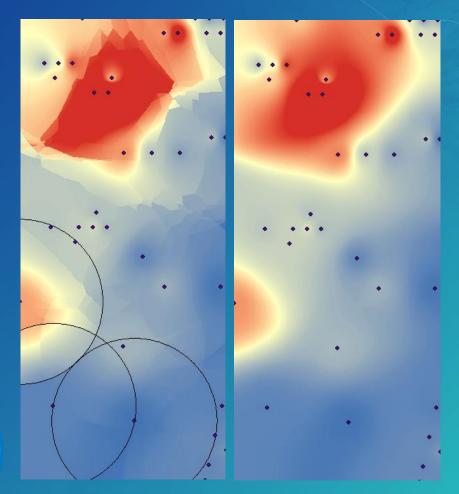
## Search neighborhood - Smooth



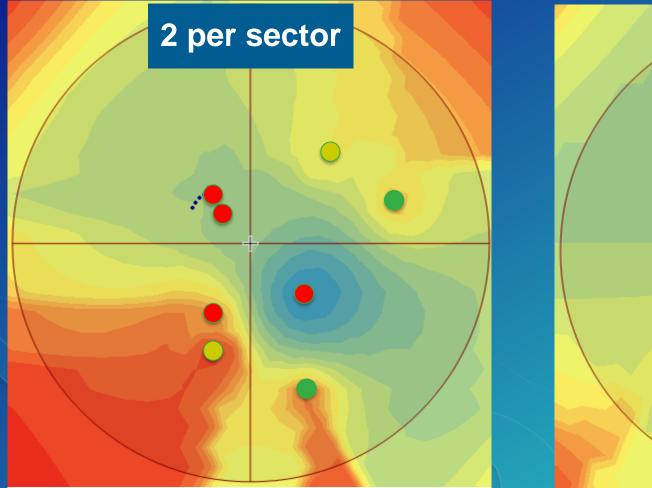
## Search neighborhood - Smooth

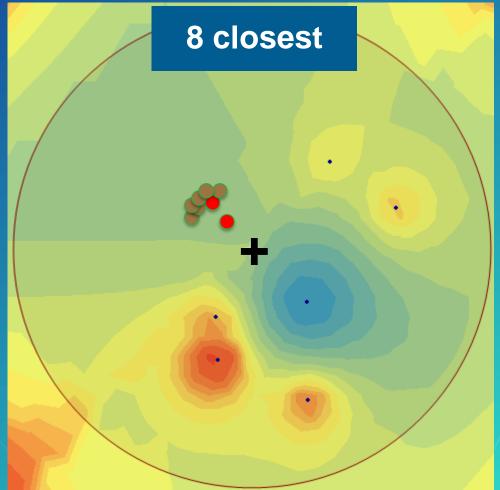


Unlike smoothing the output, this method modifies the weights



## Search neighborhood - Standard





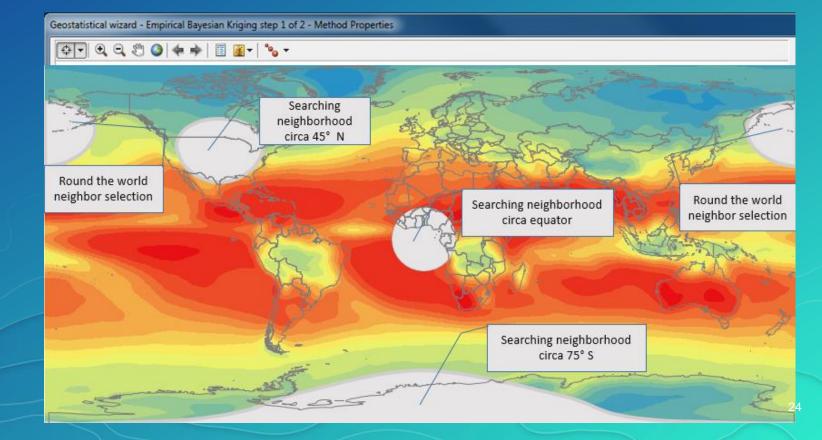
#### Chordal distances

Only for EBK and EBK Regression Prediction

- Automatically kicks in when data are in GCS
- The chordal distance between any two points is the straight-line distance that connects the two points.
- This line will go through the earth rather than along its surface.

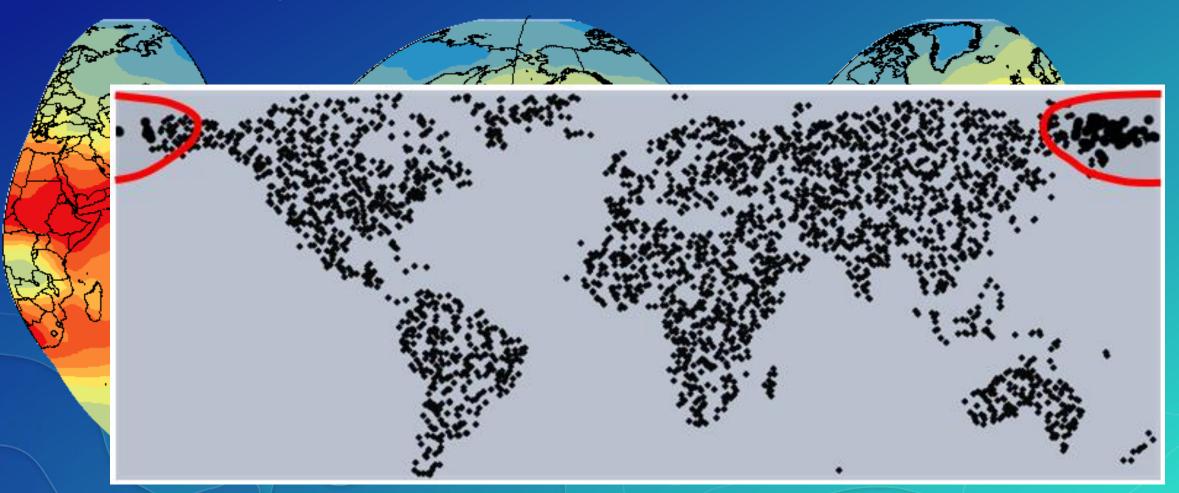
Distance between LA and New York Geodesic = 3,939.1 km Chordal = 3,877.0 km Difference = 62.1 km (1.5%)

Speed!



#### Chordal distances

Only for EBK and EBK Regression Prediction



#### Cross validation / Validation

#### Geostatistical Wizard - Kriging - Cross validation

Predicted Error Standardized Error Normal QQ Plot Distribution	Summ	ary Table					****
Measured (10 <sup>3</sup> )	Source	Measured	Predicted	Error	Standard Erro	Standardized Error	Normal Value
3.199	1	428.2999877	461.08317957	32.783191783	158.922240!	0.20628448013536!	0.4026278602534
• Count		8319		89.93730532	159.6068063	-0.56349291958548	-1.468515765765
2.748 Average CRPS		36.836434	0137905	3.7207645107	157.311614!	0.02365219199304	-0.079630867952
2.298 Inside 90 Percent Interv	-	91.104700		104.26072935	157.839014	0.66055106665327;	1.4493019584628
••				2.311307319	156.356398!	0.27060809598049;	0.564307656901(
1.847 Inside 95 Percent Interv	al	94.951316		84.05529627	159.962529:	-0.52546866281542	-1.423933648983
1.396 Mean		1.2625696	5544221	91.61338229	156.753627(	-0.58444186600423	-1.499322001875
Root-Mean-Square		77.755802	3717596	10.60162363	165.966988	0.66640736706294	1.4614664177157
0.945 Mean Standardized		0.0137622	990752729	18.36437298	153.211852	-0.11986261310177	-0.545329236077
0.494 Root-Mean-Square Star	ndardized	0.9613437	42125846	38.25203855	168.021862	0.822821724627834	1.7468766871662
Average Standard Error		82.963941	5251061	1.13604292(	162.759960(	0.43706107387348!	0.9664597708545
0.043 0.438 0.832 1.227 1.621 2.016				20.38557130	158.2113264	-0.12885026477378	-0.576361720528
Predicted (10 <sup>3</sup> )	13	486	605.71862993	119.71862993	167.743938	0.71369869361293!	1.5534026919965
• Value  — Reference Line  — Regression Line	14	412.8999938	397.1256724€	-15.77432143	151.138524	-0.10436995771118	-0.493259080077
egression function: 1.04431 * x + -34.8993	< <b>1</b>	100	140 4005 4005	10.00015300	450.000400	0.000004.40475.04	4 4 4 5 7 5 7 7 7 7 7 7 7 7 7 7 7 7 7 7

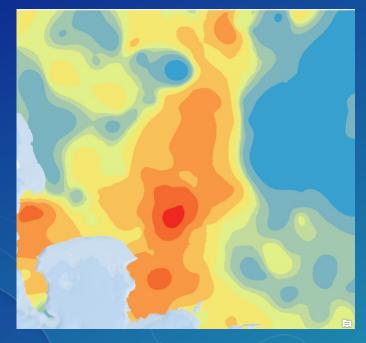
< Back Next >

Finish

26

## Output surfaces

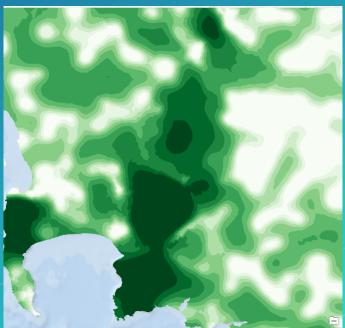
#### Prediction



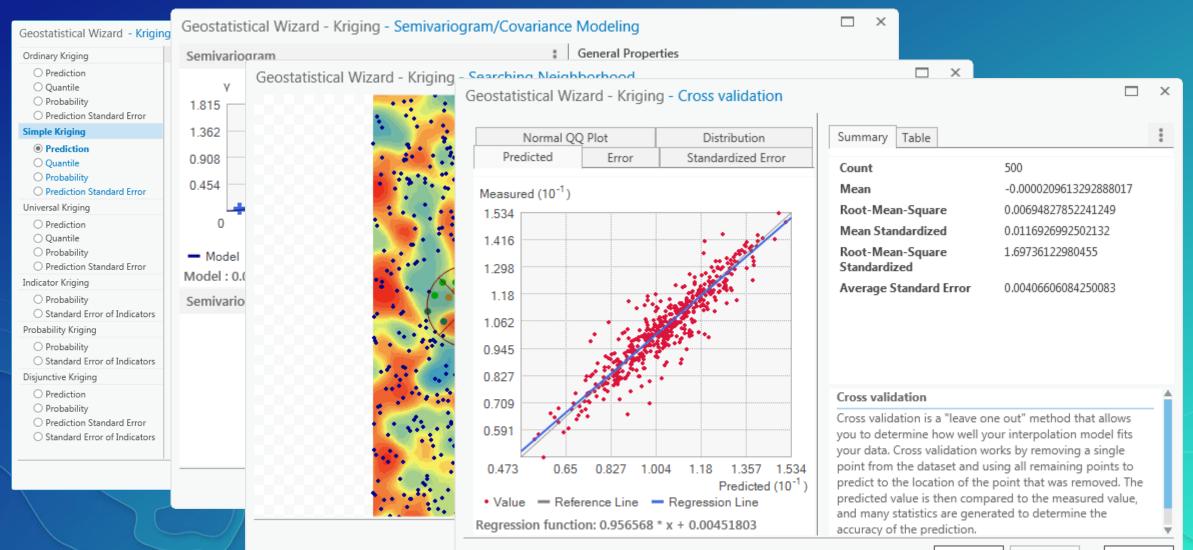
#### Standard error of prediction



#### Probability that rainfall exceeds 900mm

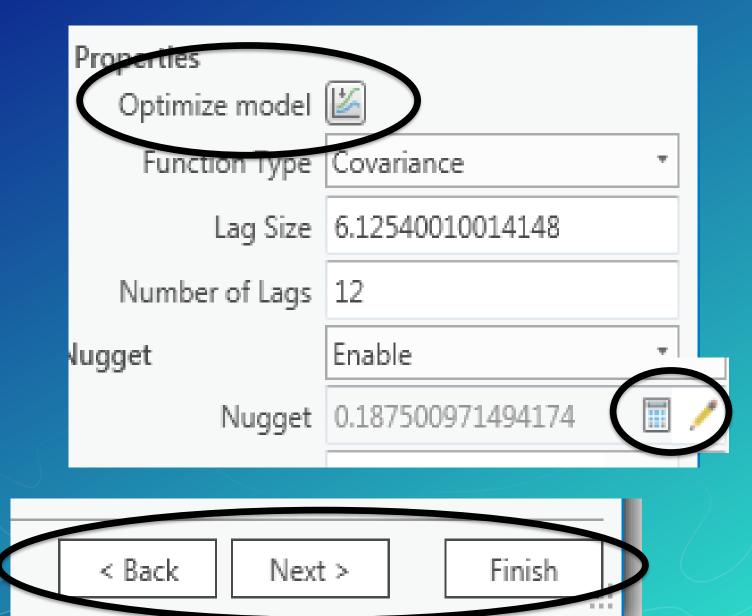


#### **Interactive Wizard**



Finish

## Interactive Wizard



#### Barriers

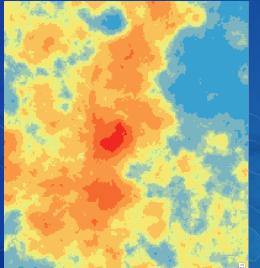


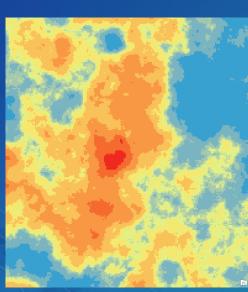
30

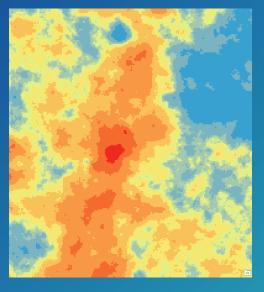
#### **Gaussian Geostatistical Simulations**

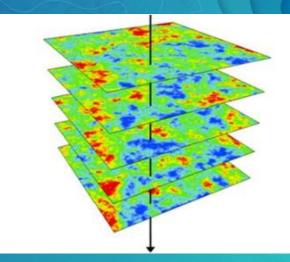
Create multiple versions (realizations) of a surface to perform risk analysis.

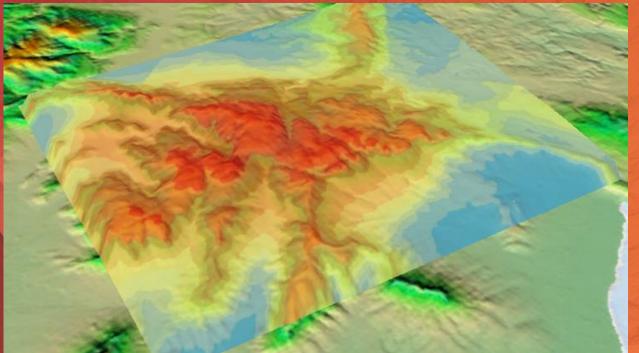
• Any realization might be the "real" thing!





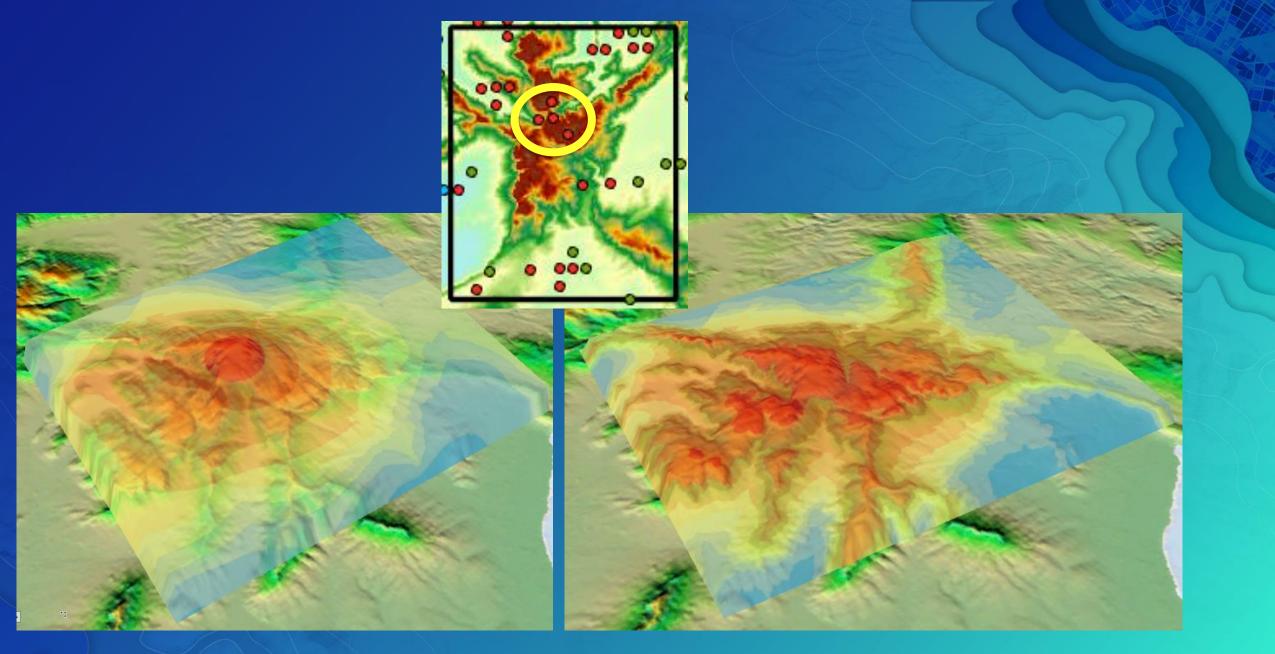




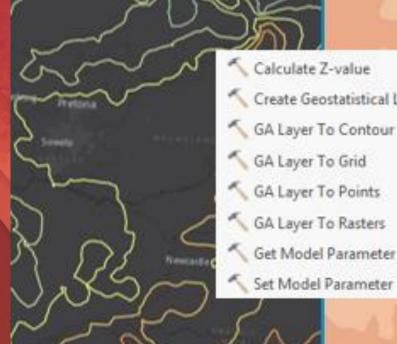


## **EBK Regression Prediction**

Eric Krause



## **EBK Regression Prediction**



Calculate Z-value 🔨 Create Geostatistical Layer 🔨 GA Layer To Contour 🔨 GA Layer To Grid 🔨 GA Layer To Points SA Layer To Rasters Get Model Parameter

Scetona

Subvito

# **Geostatistical layers**

Eric Krause

#### Where is GA used?

- Anyone who needs to statistically explore data and create surfaces for a number of variables will benefit from this statistical software package.
- Some of the various fields that use ArcGIS Geostatistical Analyst include:
  - agriculture,
  - geology,
  - meteorology,
  - hydrology,
  - archaeology,
  - forestry,
  - oceanography,
  - fishery,
  - health care, and
  - environmental studies.

## Tips & Tricks

- Use Mask when creating a raster
  - 8700 pixels inside (55,000 outside) Japan (6 1/2 times)

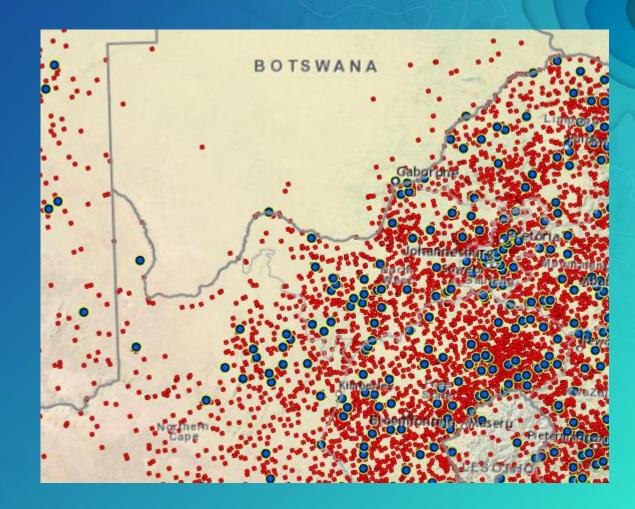
Geoprocessing	+ □ ×	
Empirical Bayes	ian Kriging 🗧	
Parameters   Environments	?	
Input features pts	Geoprocessing	×
Z value field MAP	Empirical Bayesian Kriging Parameters Environments	■ ⑦
Output geostatistical layer	Output Coordinate System	
Output raster myRaster	Geographic Transformations	- @
Output cell size 601.338979199998	Extent Default	
Data transformation type None	Snap Raster	• 🖻
Semivariogram model type Power	Cell Size Maximum of Inputs	• 🖽
> Additional Model Paramete	Mask	
<ul> <li>Search Neighborhood Para</li> <li>Output Parameters</li> </ul>	Coincident Points Mean of values at coincident locations.	
	Parallel Processing Factor	
Catalog Geoprocessing		
	Catalog Geoprocessing	Run 🕑



## Tips & Tricks

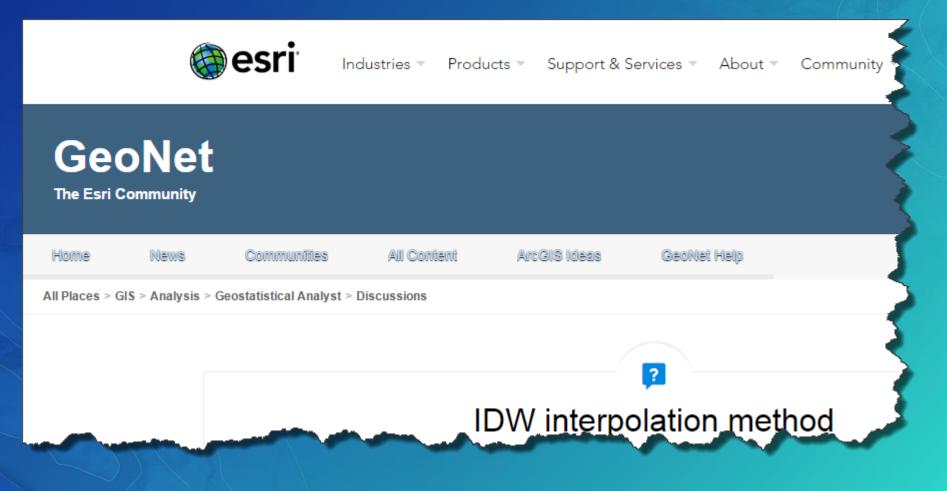
- Subset of the data
  - SubsetFeatures GP tool
  - Selection

Geopro	ocessing		<b>-</b> □ ×
$\odot$	Subset F	eatures	≡
Parame	ters   Environments		?
Input fe global	eatures L_10000pts		- 🖽
Output trainin	training feature class		+
Output test	test feature class		+
	training feature subset		80
	size units ENTAGE_OF_INPUT		•
		R	un 🕑
Catalog	Geoprocessing		



#### Conclusions

# https://geonet.esri.com/

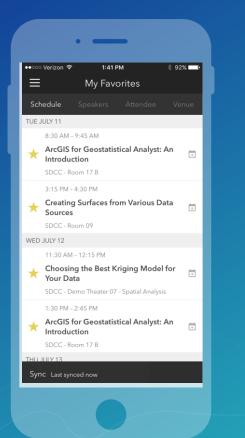


#### Please Take Our Survey on the Esri Events App!

#### Download the Esri Events app and find your event



## Select the session you attended



## Scroll down to find the survey

· —	
••∞ Verizon 🗢 1:41 PM 🕸 92% 📼 < 🖆	
ArcGIS for Geostatistical Analyst: An Introduction SDCC - Room 17 B	
🛗 Add Reminder	
TIME	
★ Jul 11, 8:30 AM - 9:45 AM	
ADDITIONAL OFFERINGS	
Jul 12, 1:30 PM - 2:45 PM >	
DESCRIPTION	
This session will present the features of the Geostatistical Analyst extension and will focus on how to use the available tools and techniques to analyze and interpolate spatial data. Performing a valid statistical analysis is a several step process which includes data exploration, choosing a model, building and validating the model, and interpreting the results. Data exploration can be done with the	
27 Dews & Die	

#### Complete Answers and Select "Submit"

●●○○○ Verizon 🗢	1:42 PM	∦ 92%	Ð
< ArcGI	S for Geostatis	stical A 📋	
Session Categorie		>	
- Technical Wo	rkshops (454)		
Session Categorie	S	>	
Performing An	alysis (103)		
FEEDBACK			
Title and Descri	ption Consistent w	ith Content	
Low 1 2		5) High	
Well Organized	/Clear Presentatior	ı	
Low 1 2		5 High	
Public Speaking	Skills		
Low 1 2	4(	5 High	
The content of t work	he workshop was i	relevant to my	
No			
The workshop p	provided information	on or techniques l	
ີ News &	Info		



#### Find "optimal" script

```
iterationsList = [30, 100, 500, 1000]
subsetSizeList = [100, 200, 500]
overlapList = [1,3,5]
semivariogram List = [['POWER', 'NONE'], ['LINEAR', 'NONE'], ['THIN PLATE SPLINE', 'NONE'],
                       ['EXPONENTIAL DETRENDED', 'EMPIRICAL'], ['EXPONENTIAL', 'EMPIRICAL'],
                       ['K BESSEL DETRENDED', 'EMPIRICAL'], ['K BESSEL', 'EMPIRICAL'],
                       ['WHITTLE', 'EMPIRICAL'], ['WHITTLE DETRENDED', 'EMPIRICAL']]
try:
    for iterations in iterationsList:
        for i in range(0, len(semivariogram List)):
            for subsetSize in subsetSizeList:
                for overlap in overlapList:
                    svg = semivariogram List[i][0]
                    transf=semivariogram List[i][1]
                    t1 = time.time()
                    result = arcpy.EmpiricalBayesianKriging ga(inFC, inField, outLyr, outRas, cellsize,
                                                                transf, subsetSize, overlap, iterations,
                                                                sn, 'PREDICTION', '', '', svg)
                    t2 = time.time()
                    tebk = t2-t1
                    cv = arcpy.CrossValidation ga(outLyr)
                    rmseValue = cv.rootMeanSquare
                    rmsStd = cv.rootMeanSquareStandardized
                    toWriteList = [rmseValue, rmsStd, svg, transf, tebk, subsetSize, overlap, iterations]
                    outFile.writerow(toWriteList)
                    arcpy.Delete management(outLyr)
except:
    print ('Tool execution FAILED')
```

```
print (arcpy.GetMessages())
```