precipitation for your state. Derived from regression equations, the PRISM map is essentially a model and may contain unusual precipitation amounts.

6. Given a sample size of 12, illustrate the difference between a sampling method that uses the closest points and a quadrant sampling method.
7. Describe how cell densities are derived using the kernel density estimation method.
8. The power k in inverse distance weighted interpolation determines the rate of change in values from the sample points. Can you think of a spatial phenomenon that should be interpolated with a k value of 2 or higher?
9. Describe how the semivariance can be used to quantify the spatial dependence in a data set.
10. Binning is a process for creating a usable semivariogram from empirical data. Describe how binning is performed.
11. A semivariogram must be fitted with a mathematical model before it can be used in kriging. Why?
12. Both IDW and kriging use weights in estimating an unknown value. Describe the difference between the two interpolation methods in terms of derivation of the weights.
13. Explain the main difference between ordinary kriging and universal kriging.
14. The root mean square (RMS) statistic is commonly used for selecting an optimal interpolation method. What does the RMS statistic measure?
15. Explain how one can use the validation technique for comparing different interpolation methods.

APPLICATIONS: SPATIAL INTERPOLATION

This applications section has five tasks. Task 1 covers trend surface analysis. Task 2 deals with kernel density estimation. Task 3 uses IDW for local interpolation. Tasks 4 and 5 cover kriging: Task 4 uses ordinary kriging and Task 5 universal kriging. Except for Task 2, you will run spatial interpolation in Geostatistical Analyst so that you can use the cross-validation statistics such as the root mean square (RMS) statistic to compare models. Geostatistical Analyst also provides more information and a better user interface than Spatial Analyst or ArcToolbox for spatial interpolation.

Task 1: Use Trend Surface Model for Interpolation

What you need: *stations.shp,* a shapefile containing 175 weather stations in and around Idaho; and *idoutlgd,* an Idaho outline raster.

In Task 1 you will first explore the average annual precipitation data in *stations.shp,* before running a trend surface analysis.

1. Start ArcCatalog, and connect to the Chapter 15 database. Launch ArcMap. Add *stations.shp* and *idoutlgd* to Layers and rename the data frame Task 1. Make sure that both the Geostatistical Analyst and Spatial Analyst extensions are checked in the Tools menu and their toolbars are checked in the View menu.
2. Click the Geostatistical Analyst dropdown arrow, point to Explore Data, and select Trend Analysis. At the bottom of the Trend Analysis dialog, click the dropdown arrow to select *stations* for the layer and ANN_PREC for the attribute.
3. Maximize the Trend Analysis dialog. The 3-D diagram shows two trend projections: The *YZ*

plane dips from north to south, and the *XZ* plane dips initially from west to east and then rises slightly. The north–south trend is much stronger than the east–west trend, suggesting that the general precipitation pattern in Idaho decreases from north to south. Close the dialog.

4. Click the Geostatistical Analyst dropdown arrow, and select Geostatistical Wizard. The opening panel lets you choose the input data and geostatistical method. Click the Input Data dropdown arrow and select *stations*. Click the Attribute dropdown arrow and select ANN_PREC. In the Methods frame, click Global Polynomial Interpolation. Click Next.
5. The Step 1 panel lets you choose the power of the trend surface model. The Power list provides the choice from 1 to 10. Select 1 for the power. The next panel shows scatter plots (Predicted versus Measured values, and Error versus Measured values) and statistics related to the first-order trend surface model. The RMS statistic measures the overall fit of the trend surface model. In this case, it has a value of 6.073. Click Back and change the power to 2. The RMS statistic for the power of 2 has a value of 6.085. Repeat the same procedure with other power numbers. The trend surface model with the lowest RMS statistic is the best overall model for this task. For ANN_PREC, the best overall model has the power of 5. Change the power to 5, and click Finish. Click OK in the Method Summary dialog.

Q1. What is the RMS statistic for the power of 5?

6. *Global Polynomial Interpolation Prediction Map* is a Geostatistical Analyst (ga) output layer and has the same area extent as *stations*. Right-click *Global Polynomial Interpolation Prediction Map* and select Properties. The Symbology tab has four Show options: Hillshade, Contours, Grid, and Filled Contours. Uncheck all Show boxes except Filled Contours, click Filled Contours, and then click on Classify. In the Classification dialog, select the Manual method and 7 classes. Then enter the class breaks of 10, 15, 20, 25, 30, and 35 between the Min and Max values. Click OK to dismiss the dialogs. The contour (isohyet) lines are color-coded.
7. To clip *Global Polynomial Interpolation Prediction Map* to fit Idaho, first convert the ga data set to a raster. Right-click *Global Polynomial Interpolation Prediction Map*, point to Data, and select Export to Raster. In the Export to Raster dialog, enter 2000 (meters) for the cell size and specify *trend5_temp* for the output raster. Click OK to export the data set. (The GA Layer to Grid tool in Geostatistical Analyst Tools can also perform the conversion.) Add *trend5_temp* to the map. (Extreme cell values in *trend5_temp* are located outside the state border.)
8. Now you are ready to clip *trend5_temp*. Double-click the Extract by Mask tool in the Spatial Analyst Tools/Extraction toolset. In the next dialog, select *trend5_temp* for the input raster, select *idoutlgd* for the input mask data, specify *trend5* for the output raster, and click OK. *trend5* is the clipped *trend5_temp*.
9. You can generate contours from *trend5* for data visualization. Click the Spatial Analyst dropdown arrow, point to Surface Analysis, and select Contour. In the Contour dialog, make sure that *trend5* is the input surface, enter 5 for the contour interval and 10 for the base contour, and specify *trend5ctour.shp* for the output features. Click OK. To label the contour lines, right-click *trend5ctour* and select Properties. On the Labels tab, check the box to label features in this layer, select CONTOUR from the Label Field dropdown list, and click OK. The map now shows contour labels.

Task 2: Use Kernel Density Estimation Method

What you need: *deer.shp,* a point shapefile showing deer locations.

Task 2 uses the kernel density estimation method to compute the average number of deer sightings per hectare from *deer.shp.* Deer location data have a 50-meter minimum discernible distance; therefore, some locations have multiple sightings.

1. Insert a new data frame in ArcMap and rename it Task 2. Add *deer.shp* to Task 2. Select Properties from the context menu of *deer.* On the Symbology tab, select Quantities and Graduated symbols in the Show box and select SIGHTINGS from the Value dropdown list. Click OK. The map shows deer sightings at each location in graduated symbols.

Q2. What is the value range of SIGHTINGS?

2. Click Show/Hide ArcToolbox Window to open ArcToolbox. Double-click the Kernel Density tool in the Spatial Analyst Tools/Density toolset. Select *deer* for the input point features, select SIGHTINGS for the population field, specify *kernel_d* for the output raster, enter 100 for the output cell size, enter 100 for the search radius, and select HECTARES for the area units. Click OK to run the command. *kernel_d* shows deer sighting densities computed by the kernel density estimation method.

Q3. What is the value range of deer sighting densities?

3. To view *kernel_d* on top of *deer,* you can use the transparent option. Right-click *kernel_d,* select Properties, and click Yes to build unique histogram. On the Display tab, enter 30% transparency. You can now see two layers superimposed on top of one another.

Task 3: Use IDW for Interpolation

What you need: *stations.shp* and *idoutlgd,* same as in Task 1.

This task lets you create a precipitation raster using the IDW method.

1. Insert a new data frame in ArcMap and rename it Task 3. Add *stations.shp* and *idoutlgd* to Task 3.
2. Click the Geostatistical Analyst dropdown arrow and select Geostatistical Wizard. Select *stations* for the input data and ANN_PREC for the attribute. Click Inverse Distance Weighting in the Methods frame. Click Next.
3. The Step 1 panel includes a graphic frame and a method frame for specifying IDW parameters. The default IDW method uses a power of 2, 15 neighbors (control points), and a circular area from which control points are selected. The graphic frame shows *stations* and the points and their weights (shown in percentages and color symbols) used in deriving the estimated value for a test location. You can use the Identify Value tool to click any point within the graphic frame and see how the point's estimated value is derived.
4. The Optimize Power Value button is included in the Step 1 panel. Because a change of the power value will change the estimated value at a point location, you can click the button and ask Geostatistical Wizard to find the optimal power value while holding other parameter values constant. Geostatistical Wizard employs the cross-validation technique to find the optimal power value. Click the Optimize Power Value button, and the Power field shows a value of 3.191. Click Next.
5. The Step 2 panel lets you examine the cross-validation results including the RMS statistic.

Q4. What is the RMS statistic when you use the default parameters including the optimal power value?

Q5. Change the power to 2 and the number of neighbors to include to 10 (including at least 6). What RMS statistic do you get?

6. Set the parameters back to the default including the optimal power value. Click Finish. Click OK in the Method Summary dialog. You can follow the same steps as in Task 1 to convert *Inverse Distance Weighting Prediction Map* to a raster, to clip the raster by using *idoutlgd* as the analysis mask, and to create isolines from the clipped raster.

Task 4: Use Ordinary Kriging for Interpolation

What you need: *stations.shp* and *idoutlgd.*

In Task 4, you will first examine the semivariogram cloud from 175 points in *stations.shp*. Then you will run ordinary kriging on *stations.shp* to generate an interpolated precipitation raster and a standard error raster.

1. Select Data Frame from the Insert menu in ArcMap. Rename the new data frame Tasks 4&5, and add *stations.shp* and *idoutlgd* to Tasks 4&5. First explore the semivariogram cloud. Click the Geostatistical Analyst dropdown arrow, point to Explore Data, and select Semivariogram/Covariance Cloud. Select *stations* for the layer and ANN_PREC for the attribute. To view all possible pairs of control points in the cloud, enter 82,000 for the lag size and 12 for the number of lags. Use the mouse pointer to drag a box around the point to the far right of the cloud. Check *stations* in the ArcMap window. The highlighted pair consists of the control points that are farthest apart in *stations*. The semivariogram shows a typical pattern of spatially correlated data: the semivariance increases rapidly to a distance of about 200,000 meters (2.00×10^5) and then gradually decreases.

2. To zoom in on the distance range of 200,000 meters, change the lag size to 10,000 and the number of lags to 20. The semivariance actually starts to level off at about 125,000 meters. To see if the semivariance has the directional influence, check the box to show search direction. You can change the search direction by either entering the angle direction or using the direction controller in the graphic. Drag the direction controller in the counterclockwise direction from 0° to 180° but pause at different angles to check the semivariogram. The fluctuation in the semivariance tends to increase from northwest (315°) to southwest (225°). This indicates that the semivariance has the directional influence. Close the Semivaraince/Covariance Cloud window. Clear selected features.

3. Select Geostatistical Wizard from the Geostatistical Analyst menu. Select *stations* for the input data and ANN_PREC for the attribute. Click Kriging in the Methods frame. Click Next. The Step 1 panel lets you select the kriging method. Select Ordinary Kriging/Prediction Map. Click Next.

4. The Step 2 panel shows the semivariogram/covariance view, which is similar to the semivariogram/covariance cloud except that the semivariance data have been averaged by distance and direction (i.e., binned). The Models frame lets you choose a mathematical model to fit the empirical semivariogram. First change the lag size to 40,000 and the number of lags to 12. A general guideline in choosing the lag size and number of lags is that their product should be about half the longest distance among all pairs of control points. The longest distance in *stations* is slightly over 960,000 meters. Click on Exponential and then check the box for Anisotropy. Geostatistical Analyst automatically calculates the optimal angle direction for anisotropy. Check the box to show search direction in the semivariogram/covariance surface view, enter a bandwidth of 6, and move the direction controller to see if the calculated angle direction yields the best fit between the model (represented by the purple line) and the empirical semivariogram. Click Next.

5. The Step 3 panel lets you choose the number of neighbors (control points), and the sampling method. Click Next.

6. The Step 4 panel shows the cross-validation results. The Chart frame offers four types of scatter plots (Predicted versus Measured values, Error versus Measured values, Standardized Error versus Measured values, and Quantile-Quantile plot for Standardized

Error against Normal values). The Prediction Errors frame lists cross-validation statistics, including the RMS statistic.

Q6. What is the RMS value?

Q7. Try a different mathematical model (e.g., Spherical, Gaussian, etc.) and see if a better set of cross-validation statistics than the Exponential model is obtainable.

7. Click Finish in the Step 4 panel. Click OK in the Method Summary dialog. *Ordinary Kriging Prediction Map* is added to the map. To derive a prediction standard error map, you will click Ordinary Kriging/Prediction Standard Error Map in the Step 1 panel and repeat panels 2 to 4.
8. You can follow the same steps as in Task 1 to convert *Ordinary Kriging Prediction Map* and *Ordinary Kriging Prediction Standard Error Map* to rasters, to clip the rasters by using *idoutlgd* as the analysis mask, and to create isolines from the clipped rasters.

Task 5: Use Universal Kriging for Interpolation

What you need: *stations.shp* and *idoutlgd.*

In Task 5 you will run universal kriging on *stations.shp.* The trend to be removed from the kriging process is the first-order trend surface.

1. Click the Geostatistical Analyst dropdown arrow and select Geostatistical Wizard. Select *stations* for the input data and ANN_PREC for the attribute. Click Kriging in the Methods frame. Click Next.
2. In the Step 1 panel, click Universal Kriging/Prediction Map in the Geostatistical Methods frame. Select First from the Order of Trend dropdown list. Click Next.
3. The Step 2 panel shows the first-order trend that will be removed from the kriging process. Click Next.
4. In the Step 3 panel, enter 10,000 for the lag size and 15 for the number of lags. Click on Spherical and then check the box for Anisotropy. Click Next.
5. Take the default values for the number of neighbors and the sampling method. Click Next.
6. The Step 5 panel shows the cross-validation results. Although the RMS value is about the same as ordinary kriging in Task 4, the standardized RMS value is lower than ordinary kriging. This means that the estimated standard error from universal kriging is not as reliable as that from ordinary kriging.

Q8. What is the standardized RMS value from the Step 5 panel?

7. Click Finish in the Step 5 panel. Click OK in the Method Summary dialog. *Universal Kriging Prediction Map* is an interpolated map from universal kriging. To derive a prediction standard error map, you will click Universal Kriging/Prediction Standard Error Map in the Step 1 panel and repeat panels 2 to 5.
8. You can follow the same steps as in Task 1 to convert *Universal Kriging Prediction Map* and *Universal Kriging Prediction Standard Error Map* to rasters, to clip the rasters by using *idoutlgd* as the analysis mask, and to create isolines from the clipped rasters.

Challenge Task

What you need: *stations.shp* and *idoutlgd.*

This challenge task asks you to compare the interpolation results from two spline methods in Geostatistical Analyst. Except for the interpolation method, you will use the default values for the challenge task. The task has three parts: one, create an interpolated raster using regularized splines; two, create an interpolated raster using thin-plate splines with tension; and three, use a local operation to compare the two rasters. The result can show the difference between the two interpolation methods.

1. Create a Radial Basis Functions Prediction map by using the kernel function of

Completely Regularized Spline. Convert the map to a raster, and save the raster as *regularized* with a cell size of 2000.

2. Create a Radial Basis Functions Prediction map by using the kernel function of Spline with Tension. Convert the map to a raster, and save the raster as *tension* with a cell size of 2000.
3. Select Options from the Spatial Analyst menu. On the General tab, select *idoutlgd* for the analysis mask.
4. Use Raster Calculator from the Spatial Analyst menu to subtract *tension* from *regularized*.
5. The calculation result shows the difference in cell values between the two rasters within *idoutlgd*. Display the difference raster in three classes: lowest value to −0.5, −0.5 to 0.5, and 0.5 to highest value.

Q1. What is the range of cell values in the difference raster?

Q2. What does a positive cell value in the difference raster mean?

Q3. Is there a pattern in terms of where high cell values, either positive or negative, are distributed in the difference raster?

REFERENCES

Ackerman, W. V., and A. T. Murray. 2004. Assessing Spatial Patterns of Crime in Lima, Ohio. *Cities* 21: 423–37.

Bailey, T. C., and A. C. Gatrell. 1995. *Interactive Spatial Data Analysis.* Harlow, England: Longman Scientific & Technical.

Baxter, S. J., and M. A. Oliver. 2005. The Spatial Prediction of Soil Mineral N and Potentially Available N Using Elevation. *Geoderma* 128: 325–39.

Beguería, S., and S. M. Vicente-Serrano. 2006. Mapping the Hazard of Extreme Rainfall by Peaks over Threshold Extreme Value Analysis and Spatial Regression Techniques. *Journal of Applied Meteorology and Climatology* 45: 108–24.

Burrough, P. A., and R. A. McDonnell. 1998. *Principles of Geographical Information Systems.* Oxford, England: Oxford University Press.

Cai, Q., G. Rushton, B. Bhaduri, E. Bright, and P. Coleman. 2006. Estimating Small-Area Populations by Age and Sex Using Spatial Interpolation and Statistical Inference Methods. *Transactions in GIS* 10: 577–98.

Calvo, E., and M. Escolar. 2003. The Local Voter: A Geographically Weighted Approach to Ecological Inference. *American Journal of Political Science* 47: 189–204.

Carroll, S. S., and N. Cressie. 1996. A Comparison of Geostatistical Methodologies Used to Estimate Snow Water Equivalent. *Water Resources Bulletin* 32: 267–78.

Chang, K., and Z. Li. 2000. Modeling Snow Accumulation with a Geographic Information System. *International Journal of Geographical Information Science* 14: 693–707.

Chung, K., D. Yang, and R. Bell. 2004. Health and GIS: Toward Spatial Statistical Analyses. *Journal of Medical Systems* 28: 349–60.

Cressie, N. 1991. *Statistics for Spatial Data.* Chichester, England: Wiley.

Daly, C., R. P. Neilson, and D. L. Phillips. 1994. A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology* 33: 140–58.

Davis, J. C. 1986. *Statistics and Data Analysis in Geology,* 2d ed. New York: Wiley.

Declercq, F. A. N. 1996. Interpolation Methods for Scattered Sample Data: Accuracy, Spatial Patterns, Processing Time. *Cartography and Geographic Information Science* 23: 128–44.

Eriksson, M., and P. P. Siska. 2000. Understanding Anisotropy Computations. *Mathematical Geology* 32: 683–700.

Foody, G. M. 2004. Spatial Nonstationarity and Scale-Dependency in the Relationship Between Species Richness and Environmental Determinants for the Sub-Saharan Endemic Avifauna. *Global Ecology and Biogeography* 13: 315–20.