THE EFFECTS OF INFLUENTIAL OBSERVATIONS ON SAMPLE SEMIVARIOGRAMS

by

Sabyasachi Basu and Richard F. Gunst
Department of Statistical Science
Southern Methodist University
Dallas, TX 75275-0332, USA

Technical Report No. SMU/DS/TR-270

June 1994

THE EFFECTS OF INFLUENTIAL OBSERVATIONS ON SAMPLE SEMIVARIOGRAMS

Sabyasachi Basu and Richard F. Gunst

Department of Statistical Science, Southern Methodist University,

Dallas, Texas 75275-0332, USA

ABSTRACT

Optimal estimation of spatial characteristics such as the values of regionalized variables or the means of random fields is often accomplished using kriging methods. These methods rely on satisfactory estimation of spatial semivariograms and the fitting of semivariogram models. One of the many computational challenges to the fitting of semivariogram models is the detection and accommodation of influential observations. Influential observations can have a dramatic effect on sample semivariograms because each observation is used many times in the calculation of the semivariogram values. In this paper methods for detecting influential observations are discussed. Readily computable diagnostics for generalized least squares estimators are applied to kriging model fits and are shown to be highly effective for identifying influential observations. An analysis of regional temperature anomalies demonstrates that deletion of influential observations may be required in order to obtain satisfactory semivariogram model fits.

Key words and Phrases: diagnostics; kriging; optimal interpolation; regression; robust estimation; spatial modeling.

AMS subject classifications: 62M30, 62M40, 86A32

1. Introduction

The intense contemporary concern over environmental issues is one of the motivations for the current emphasis on spatial statistical modeling of data collected regionally and around the globe. Spatial statistical models account for differences in correlations between observations taken relatively close to one another and those taken greater distances apart. One important class of spatial modeling techniques is known variously as kriging, optimal spatial averaging/interpolation, or objective analysis. Kriging consists of two basic steps: semivariogram or covariance model estimation and minimum variance linear unbiased prediction (Journel and Huijbregts, 1978; Cressie, 1991). If the semivariogram model is known or has been satisfactorily modeled, the various types of kriging and co-kriging can be performed using generalized least squares estimators of appropriately specified trend models (Stein and Corsten, 1991). However, a large number of issues remain to be resolved in the estimation of semivariogram models. Concern in this paper is with the estimation of semivariogram models when one or more of the observations may be influential.

Figure 1(a) is a plot of November 1990 sample semivariogram values calculated from temperature anomalies for European (40°N to 60°N latitude, 10°W to 40°E longitude) temperature reporting stations. Each of these stations has at least 15 years of complete monthly data over the reference period 1951-1980. Anomalies are deviations of individual monthly station temperatures from the average monthly station temperature over this reference period (Gunst, Basu, and Brunell, 1993). Anomalies, rather than actual temperatures, are commonly used to study climate changes in order to reduce local and regional effects on temperature trends (Hansen and Lebedeff, 1987; Jones, Raper,

Bradley, Diaz, Kelly, and Wigley 1986). The calculations used to obtain the sample semivariogram values in Figure 1(a) will be detailed in the next section.

Of interest here is the occurrence of spikes in the plotted values for the complete data set (all the European stations having at least 15 years of complete data over the reference period). The spikes are important because they greatly affect the fitting of a smooth nonlinear semivariogram model to the semivariogram values. It is common for such attempts to fail because of lack of convergence, as occurs with the fitting of a Gaussian semivariogram model to these values. Spikes such as these also affect the choice of a semivariogram model because they have a great influence on the estimation of semivariogram model parameters and thus to the adequacy of the fit.

Computational difficulties in fitting the sample semivariogram values to the complete data in Figure 1(a) led to an investigation of various contour plots and other graphical displays of the data, three of which are shown in Figure 1. The contours in Figure 1(b) sharply increase around the Copenhagen station (latitude 55.7^{o}N, longitude 12.6^{o}E). The perspective plot (S-Plus, 1991) in Figure 1(c) clearly shows the discrepant value for this anomaly, as does the box plot in Figure 1(d). The November 1990 Copenhagen anomaly is not only discrepant from the anomalies for all stations in its vicinity, it is also discrepant from the other November anomalies for this station over the last century. The November 1990 Copenhagen temperature anomaly is 5°C. Over the previous 200 years the next warmest November anomaly for Copenhagen is 2.6^{o}C. The quartiles of November anomalies for Copenhagen are -2.5^{o}C, -1.1^{o}C, and 0°C with a mean of -1.23^{o}C.

Elimination of the November 1990 Copenhagen anomaly results in the smoother semivariogram plot shown in Figure 1(a), convergence of the nonlinear algorithm for

fitting a Gaussian semivariogram model to the semivariogram values, and more realistic estimates of the model parameters. It is important to note that this one observation affected several of the sample semivariogram values in Figure 1(a), as indicated by the absence of several of the severest spikes in the semivariogram plot when Copenhagen is deleted. While deletion of the discrepant Copenhagen anomaly is not the only form of accommodation that can be taken, one could argue that it is warranted in this case because of the severe effects the anomaly has on the sample semivariogram values and on the fitted semivariogram models, and that doing so provides a better representation of the spatial variability of European temperature anomalies for November 1990.

When spatially modeling large data sets, it may be prohibitively laborious to critically examine all data values in order to determine whether there may be influential observations present. For example, the temperature anomaly data set that provides the focal point of this research consists of monthly anomalies for almost 1,900 temperature stations that have various lengths of temperature records from the mid 1800s to 1990. Nevertheless, the problems noted above necessitate that influential observations be identified and a decision be made regarding whether to retain or otherwise accommodate aberrant data values. Robust estimation of the semivariogram values can reduce the effects of aberrant observations and rectify the computational difficulties associated with variogram model fitting. Alternatively, the detection and deletion of severely influential observations might suffice to ameliorate the computational and estimation difficulties.

Because of the lack of statistical independence of the anomaly differences used in the estimation of semivariograms, it is not clear whether robust procedures or the identification and elimination of aberrant data values would be more advantageous. While the context of the problem and characteristics of a data set should always be used to help

determine how to accommodate influential observations, influence diagnostics should be included in any such investigation. In this paper both influence diagnostics and robust estimation are studied with regard to their impact on semivariogram estimation and modeling.

2. Robust Semivariogram Estimators

A spatial random function is a random variable defined over a continuous spatial domain D: $\{Z(\mathbf{s}), \mathbf{s} \in D\}$, where for fixed $\mathbf{s}$, $Z(\mathbf{s})$ is a random variable. A realization of this random function $\{z(\mathbf{s}_i), \mathbf{s}_i \in D\}$ is often termed a regionalized variable. The theoretical semivariogram associated with a random function is defined as

$$\gamma(\mathbf{s}_1, \mathbf{s}_2) = \mathrm{var}\{Z(\mathbf{s}_1) - Z(\mathbf{s}_2)\}/2 \ .$$

In this paper, discussion is restricted to spatial second-order stationary random functions. A further simplification assumed in this paper is that the theoretical semivariogram is isotropic, not dependent on direction. Under these assumptions, the semivariogram is a function of only the distance between two spatial locations. For the fitting of temperature anomalies, great circle distances are used. Alternative methods for calculating and fitting semivariograms are available when either the stationarity or the isotropy assumptions is violated, see Journel and Huijbregts (1978) or Cressie (1991). In addition, Gunst (1994) discusses the advantages of trend removal for the reduction or elimination of anisotropy.

Two critical steps in the use of semivariograms for spatial modeling are the estimation of the semivariogram values and the fitting of these values by semivariogram models. Semivariogram models are fit to the sample semivariogram values in order to ensure that the matrix of fitted semivariogram values is conditionally negative definite, which is equivalent to the requirement that the estimated covariance matrix of predicted

spatial variates is positive definite. This property cannot be ensured simply by estimating the semivariogram values themselves. In addition, a valuable interpretive feature of semivariogram model fitting is the decomposition of spatial variability into components corresponding to model parameters such as the nugget, the sill, and the range.

In order to estimate semivariogram values, several choices concerning the grouping of the regionalized variables must be made if locations are not equally spaced on a grid. Ordinarily, pairs of regionalized variables are binned into several equally spaced lags. The lag distances are first selected and then pairs of regionalized variables whose locations are within a specified distance and direction from one another are binned together. For isotropic semivariogram values, the direction is ignored. In Figure 1(a), lag distances were set at multiples of 100 km. Thus, all pairs of station locations that were within (0 km, 100 km], regardless of direction, were binned into the first lag, those within (100 km, 200 km] were binned into the second lag, and so forth. The lag distances, multiples of 100 km, and the number of lags used, 20, were subjectively chosen after an examination of a number of alternative choices. The choice of lag distances and number of lags are not the focus of this paper.

The traditional semivariogram estimator is

$$\hat{\gamma}(\mathbf{d}) = \sum_{N(d)} \{z(\mathbf{s}_i) - z(\mathbf{s}_j)\}^2 / 2n_d \ , \tag{2.1}$$

where N(d) denotes the set of all pairs of locations binned together at nominal lag distance d and n_d is the number of such pairs of locations. For a fixed distance d, the classical estimator is an unbiased estimator of the semivariogram. The sample semivariograms plotted in Figure 1(a) were calculated using equation (2.1). Under intrinsic stationarity

assumptions, the sample semivariogram estimator (2.1) is unbiased. If a mean shift is added to the value of a regionalized variable at location $\mathbf{s}_m$, one can show that the sample semivariogram has expectation

$$E\{\hat{\gamma}(\mathbf{d})\} = \gamma(\mathbf{d}) + \frac{N_m(d)}{N(d)}\delta_m^2$$

where $N_m(d)$ is the number of location pairs at nominal lag distance d that include $\mathbf{s}_m$ and δ_m is the magnitude of the mean shift. This quantification of the effect of a mean shift can account for the spikes in the sample semivariogram plot in Figure 1(a).

Cressie and Hawkins (1980) recognized the tendency for extreme values of regionalized variables to influence the calculation of variogram values. After studying the cumulants of squared anomaly differences under normality assumptions, they recommended the following robust variogram estimator as an alternative to the traditional estimator (2.1):

$$\hat{\gamma}_{CH}(\mathbf{d}) = \frac{\left\{ \sum_{N(d)} |z(\mathbf{s}_i) - z(\mathbf{s}_j)|^{1/2} / n_d \right\}^4}{2(.457 + .494 / n_d)} . \tag{2.2}$$

Another alternative to using the traditional estimator (2.1) is to weight a function of the anomaly differences using robust estimators of location or spread. Cressie and Hawkins (1980) investigated several forms of trimming and four different M-estimators of location applied to the square roots of the anomaly differences, $y = |z(\mathbf{s}_i) - z(\mathbf{s}_j)|^{1/2}$. They reported simulation results for lag-1 semivariogram values, $\gamma(1)$, and concluded that none of these alternative robust estimators was preferable to the robust estimator (2.2). They also concluded that the robust estimator (2.2) was more efficient than the sample semivariogram (2.1). McBratney and Webster (1986), on the other hand, concluded, also

on the basis of simulation results, that neither the sample semivariogram (2.1) nor the robust semivariogram (2.2) could always be recommended over the other. Zimmerman and Zimmerman (1991) concluded from their simulation results that weighted least squares estimators of theoretical semivariogram model parameters with either the sample semivariogram (2.1) or the robust semivariogram (2.2) generally performed as well as the more computationally demanding likelihood-based methods they studied.

A further alternative to (2.1) consists of weighting the anomalies themselves. For example, weighted anomalies can be calculated as

$$z_w(\mathbf{s}_i) = m_z + w_i\{z(\mathbf{s}_i) - m_z\},$$

where m_z is a robust estimator of location and w_i is a weight assigned to the ith anomaly in a bin based on robust estimates of location and scale. The robust estimator used in the following investigation was calculated from Huber's (1973, 1981)

$$\psi(z) = \min\{cs, \max(-cs, z - m_z)\}\ , \tag{2.3}$$

where s is a robust estimate of scale and c is a "tuning" constant. The weight in $z_w(\mathbf{s}_i)$ is then $w_i = \psi\{z(\mathbf{s}_i)\}/z(\mathbf{s}_i)$. The minimum absolute deviation, MAD, was used as the estimate of scale and the tuning constant was set to 1.345. Other estimates of scale and choices for the tuning constant were investigated. The conclusions reported below do not materially change for any of the other estimates of scale or choices of the tuning constant.

Figure 2 shows the effects of applying the Cressie-Hawkins robust estimator (2.2) and weighting the anomalies using (2.3) on the sample semivariogram for the November 1990 Europe anomalies. The Cressie-Hawkins robust estimates are less affected than the traditional estimates by the Copenhagen anomaly but the spike at lag 9 (950 km) remains, as do large semivariogram values at several of the other lags. The M-estimator weighting of the individual anomalies has shifted the entire semivariogram well below the other

semivariograms. The reason for this shift is the bound placed on the magnitude of the difference $z - m_Z$ in (2.3). Anomalies that exceed this value in the final iteration of the calculation of the M-estimator are all set to the same value, $\pm 1.345s$ depending on the sign of the difference $z - m_Z$. An examination of the weights reveals that a number of anomalies in most lags are set equal to the upper or the lower bounds, not just Copenhagen. Consequently, many of the weighted anomaly differences are zero, causing the averages of the squared differences for all the lags to be negatively biased estimators of the true semivariogram values.

These features of the robust estimators can be confirmed in simulation studies. Figure 3 shows the results of a simulation of 200 realizations from a Gaussian semivariogram model

$$\gamma(\mathbf{d}) = \theta_1 + (\theta_2 - \theta_1)\{1 - \exp(-\mathbf{d}^2/\theta_3^2)\} \tag{2.4}$$

with the following parameters: nugget (θ_1) = 0.2, sill (θ_2) = 2, range (θ_3) = 7. Ninety-nine locations were randomly selected over the square region [0,20]x[0,20]. Location (10,10) was added to the 99 as the location where an influential observation would be placed. The same 100 locations were used in all replications and Euclidean distance was used in the calculation of the sample semivariogram values. Figure 3(a) does not contain an influential observation. Figure 3(b) contains an influential observation whose value was set equal to 2 at location (10,10). Figure 3(c) has an influential observation at the same location with a value equal to 4. In Figure 3(d), the value was set equal to 8. The sample semivariogram that includes all the data contains spikes and has a general tendency to overestimate the theoretical semivariogram when the influential observation is added, Figures 3(b) to 3(d). The Cressie-Hawkins robust estimator $\hat{\gamma}_{CH}$ does provide protection against catastrophic effects of the influential observation, but in doing so it is biased upward. Weighting the

anomalies shows the clear effects of the weighted observations being set equal to the upper and the lower bounds of the M-estimator weighting function. In all four plots the negative bias in the weighted semivariogram values is apparent. On the other hand, the sample semivariogram with the observation at (10,10) deleted was closest to the theoretical semivariogram in all four plots.

Other simulations were performed with different outlier locations. The results were similar to those just reported. These investigations confirm the importance of having appropriate influence diagnostics to detect the presence of unusual or discrepant regionalized variable values. Diagnostics are discussed in the next section.

3. Influence Diagnostics

One compelling characteristic of commonly used case-deletion influence diagnostics is the ability to calculate the diagnostics without refitting. Thus, a de-facto requirement for routine application of influence diagnostics is that there be computational shortcuts which enable case-deletion diagnostics to be calculated from one fit to all the data. Hoaglin and Welsch (1978) introduced this technique for least squared estimators and others have generalized the approach to a variety of models and estimators, See, for example, Christensen, Johnson, and Pearson (1992), Cook and Weisberg (1982), De Gruttola, Ware and Louis (1987), Escobar and Moser (1993), Fox, Hinkley and Larntz (1980), Pregibon (1981), and Wellman and Gunst (1991).

Several approaches were used to derive influence diagnostics in this research. One approach was to express kriging model fits as generalized least squares prediction equations and to derive regression-based influence diagnostics. Stein and Corsten (1991) showed that the various types of kriging and co-kriging model predictions could be

expressed in terms of generalized least squares predictors of the regionalized variable. This work connects the optimal linear prediction theory for linear models (e.g., Goldberger 1962) to kriging predictors. The linear model formulation readily lends itself to the derivation of influence diagnostics.

Let Σ denote the spatial covariance matrix of the spatial random variables $Z(\mathbf{s}_i)$. Suppose that one wishes to predict the value Z_0 of the regionalized variable at a location $\mathbf{s}_0$. Let σ_0 denote the spatial covariances between the $Z(\mathbf{s}_i)$ and $Z(\mathbf{s}_0)$. Let $\mathbf{z}$ denote the n-dimensional vector of observed regionalized variables. The ordinary or universal kriging predictor of Z_0 is

$$\begin{aligned}\hat{z}_0 &= \sigma_0'\Sigma^{-1}\mathbf{z} + \mathbf{x}_a'\hat{\beta} \\ &= \sigma_0'\Sigma^{-1}(\mathbf{z} - \mathbf{X}\hat{\beta}) + \mathbf{x}_0'\hat{\beta}\end{aligned}, \tag{3.1}$$

where $\hat{\beta} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{z}$, $\mathbf{X}$ is a matrix consisting of a single column of ones for ordinary kriging or a column of ones and k-1 columns of trend variables for universal kriging, and $\mathbf{x}_a = \mathbf{x}_0 - \mathbf{X}'\Sigma^{-1}\sigma_0$.

Case deletion statistics for the fit (3.1) are

$$\begin{aligned}&\hat{\beta}_{(i)} - \hat{\beta} = -c_{aa}^{-1}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_a r_a / (1 - c_{aa}^{-1}h_{aa}) \\ &D_i = h_{aa}r_a^2 / \{kc_{aa}^2(1 - c_{aa}^{-1}h_{aa})^2\} \\ &\hat{z}_i - \hat{z}_{(i)} = r_a / (1 - c_{aa}^{-1}h_{aa}) \\ \text{and}\quad &t_{a(i)} = r_a / \left\{c_{aa}^{1/2}(1 - c_{aa}^{-1}h_{aa})\right\},\end{aligned} \tag{3.2}$$

where $r_a = z_a - \mathbf{x}_a'\hat{\beta}$, $z_a = z_i - \sigma_{(i)}'\Sigma_{(ii)}^{-1}\mathbf{z}_{(i)}$, $\mathbf{x}_a = \mathbf{x}_i - \mathbf{X}_{(i)}'\Sigma_{(ii)}^{-1}\sigma_{(i)}$, $c_{aa} = \sigma_{ii} - \sigma_{(i)}'\Sigma_{(ii)}^{-1}\sigma_{(i)}$, and $h_{aa} = \mathbf{x}_a'(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{x}_a$. The first two diagnostics listed in (3.2) are generalized least squares versions of Belsley, Kuh & Welsch's DFBETA and Cook's (1977) distance statistic. They are derived in Christensen, Pearson & Johnson (1992). The third diagnostic is a generalized least squares version of Belsley, Kuh & Welsch's DFFIT and the fourth

one is a Studentized version of the adjusted residual $z_a - \mathbf{x}_a'\hat{\beta}_{(i)}$. Interestingly, the DFFIT statistic is identical to the adjusted residual because the kriging fit is an exact interpolator.

In equations (3.2), the subscript (i) denotes the deletion of the ith case. The subscript a denotes a regionalized variable or a predictor that is adjusted for other variates as indicated in the expressions following equations (3.2). The use of adjusted predictors and adjusted residuals in these diagnostics are a natural consequence of the focus of kriging on optimal prediction of regionalized variables. They are adjustments due to the assumed covariance structure of the errors in the model. Observe that the case-deletion diagnostics can be obtained from the fit to the complete data set. If, as in ordinary least squares estimation, the errors are assumed to be uncorrelated, the adjusted variates reduce to the raw predictors and residuals and the diagnostics are the usual least squares case-deletion diagnostics.

The studentized adjusted residuals $t_{a(i)}$ were calculated for the November anomalies from a constant-mean, or *ordinary*, kriging model. The values are plotted in Figure 4(a) by case number. The exceptionally large residual for the Copenhagen station is clearly evident in the plot.

Alternative diagnostics that are not regression-based case deletion diagnostics can also be derived. Computationally more demanding than the regression diagnostics are diagnostics obtained directly from nonlinear semivariogram model fits. The sensitivity of these fits to the repeated use of each observation in the sample semivariogram calculations necessitates the deletion of each observation and the recalculation of both the sample semivariogram values and the model fits. For comparison purposes, each November anomaly individually was deleted from the data set and the nonlinear least squares algorithm was run 88 times to fit the Gaussian semivariogram model to the sample

semivariograms. Case-deletion diagnostics for the estimated model parameters were calculated similar to the recommendation of Bruce and Martin (1989, equation (2.8)):

$$d_{(i)} = n(\hat{\theta}_{(i)} - \hat{\theta})'\hat{\mathbf{I}}(\hat{\theta}_{(i)} - \hat{\theta}), \tag{3.3}$$

where $\hat{\mathbf{I}}$ is the estimated information matrix from the nonlinear fit using all the anomalies and $\hat{\mathbf{I}} = n^{-1}\mathbf{F}'\mathbf{F}$, with $\mathbf{F}$ the $n \times 3$ matrix of partial derivatives of the Gaussian model evaluated at the n data values. These case-deletion diagnostics are plotted in Figure 4(b) and also clearly indicate the influence of the Copenhagen station anomaly. However, they require a great additional computational effort over the one-fit residual diagnostics derived above.

An alternative approach to the derivation of influence diagnostics is to estimate $E(y_i|y_k, k \neq i)$ and $var(y_i|y_k, k \neq i)$ under normality assumptions and then form a diagnostic based on the statistic $y_i - \hat{E}(y_i|y_k, k \neq i)$. This type of influence diagnostic and others were derived and studied but did not perform substantially different from those described above.

4. Concluding Remarks

Semivariogram model fitting is a difficult theoretical and computational problem. The effects of influential observations on semivariogram estimation and the fitting of semivariogram models are made more severe by the reuse of individual data values in and between bins. In settings such as the modeling of global temperature data, the sheer magnitude of the data files preclude the extensive hands-on exploration of subsets of the data for influential observations. Thus, diagnostics can be a valuable aid in the identification of observations that might severely affect the fitting of semivariogram models.

Acknowledgments

This research was supported by Contract DE FG03 93ER61645 from the U.S. Department of Energy.

References

Belsley, D.A., E.L. Kuh, and R.E. Welsch (1980). *Regression Diagnostics.* John Wiley, New York.

Bruce, A.G. and R.D. Martin (1989). Leave-k-out diagnostics for time series. *J. R. Statist. Soc. Ser. B* **51**, 363-424.

Christensen, R., L.M. Pearson, and W. Johnson (1992). Case-deletion diagnostics for mixed models. *Tech.* **34**, 38-45.

Cook, R.D. (1977). Detection of influential observations in linear regression. *Tech.* **19**, 15-18.

Cook, R.D. and S. Weisberg (1982). *Residuals and Influence in Regression.* Chapman and Hall, London.

Cressie, N.A.C. (1991). *Statistics for Spatial Data.* John Wiley, New York.

Cressie, N. and D.M. Hawkins (1980). Robust estimation of the variogram: I. *Math. Geol.* **12**, 115-125.

De Gruttola, V., J.H. Ware and T. Louis (1987). Influence analysis of generalized least squares estimators. *J. Amer. Statist. Assoc.* **82**, 911-917.

Escobar, L.A. and E.B. Moser (1993). A note on the updating of regression estimates. *Amer. Statist.* **47**, 192-194.

Fox, T., D. Hinkley, and K. Larntz (1980). Jackknifing in nonlinear regression. *Tech.* **22**, 29-34.

Gunst, R.F. (1994). Estimating spatial correlations from spatial-temporal meteorological data. Technical Report SMU/DS/TR-267, Department of Statistical Science, Southern Methodist University.

Gunst, R.F., S. Basu, R. and Brunell (1993). Defining and estimating global mean temperature anomalies. *J. Clim.* **6**, 1368-1374..

Hansen J. and S. Lebedeff (1987). Global trends in measured surface air temperatures. *J. Geophys. Res.* **92**, 13,345-13,372.

Hoaglin, D.C. and R.E. Welsch (1978). The hat matrix in regression and ANOVA. *Amer. Statist.* **32**, 17-22.

Huber, P. (1973). Robust regression, asymptotics, conjectures, and Monte Carlo. *Ann. Statist.* **1**, 799-821.

Huber, P. (1981). *Robust Statistics.* John Wiley, New York.

Jones, P.D., S.C.B. Raper, R.S. Bradley, H.F. Diaz, P.M. Kelly, and T.ML. Wigley (1986). Northern hemisphere surface air temperature variations: 1851-1984. *J. Clim. and Appl. Met.* **25**, 161-79.

Journel, A.G. and Ch. J. Huijbregts (1978). *Mining Geostatistics.* Academic Press, New York.

McBratney, A.B. and R. Webster (1986). Choosing functions of semivariograms of soil properties and fitting them to sampling estimates. *J. Soil Sci.* **37**, 617-639.

Pregibon, D. (1981). Logistic regression diagnostics. *Ann. Statist.* **12**, 87-100.

S-Plus (1991). *S-Plus User's Manual, Version 3.0.* Statistical Sciences, Seattle.

Stein, A. and L.C.A. Corsten (1991). Universal kriging and cokriging as a regression procedure. *Biocs.*, **47**, 575-87.

Wellman, J.M. and R. F. Gunst (1991). Influence diagnostics for linear measurement error models. *Bioka.* **78**, 373-80.

Zimmerman, D.L. and M.B. Zimmerman (1991). A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. *Tech.* 33, 77-91.

List of Figures

Figure 1. European Temperature Anomalies for November 1990.

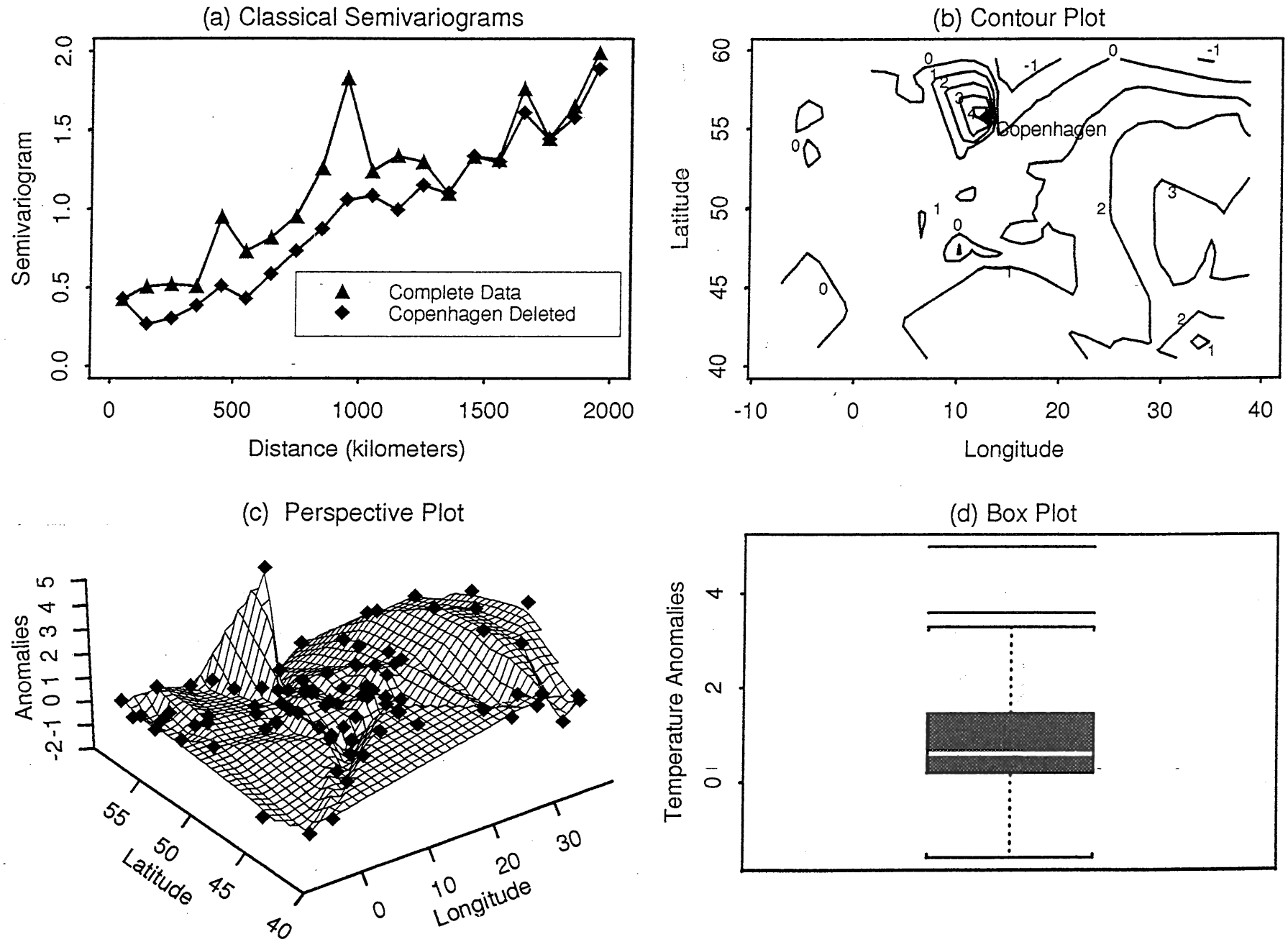


Figure 2. Sample Semivariograms for the November 1990 European Temperature Anomalies.

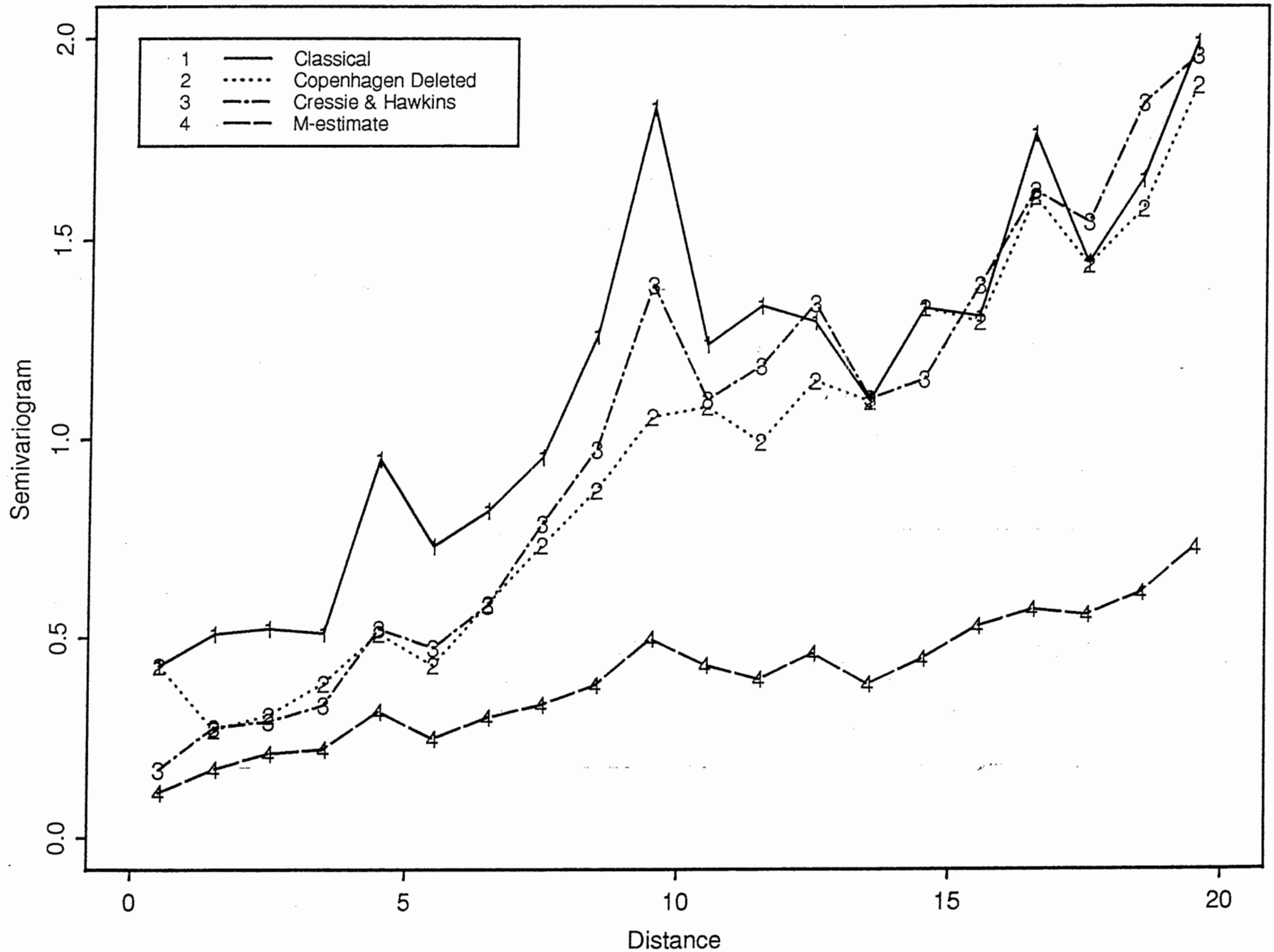


Figure 3. Simulation Averages from 200 Replications of a Gaussian Model with Nugget = 0.2 , Sill = 2 , and Range = 7 .

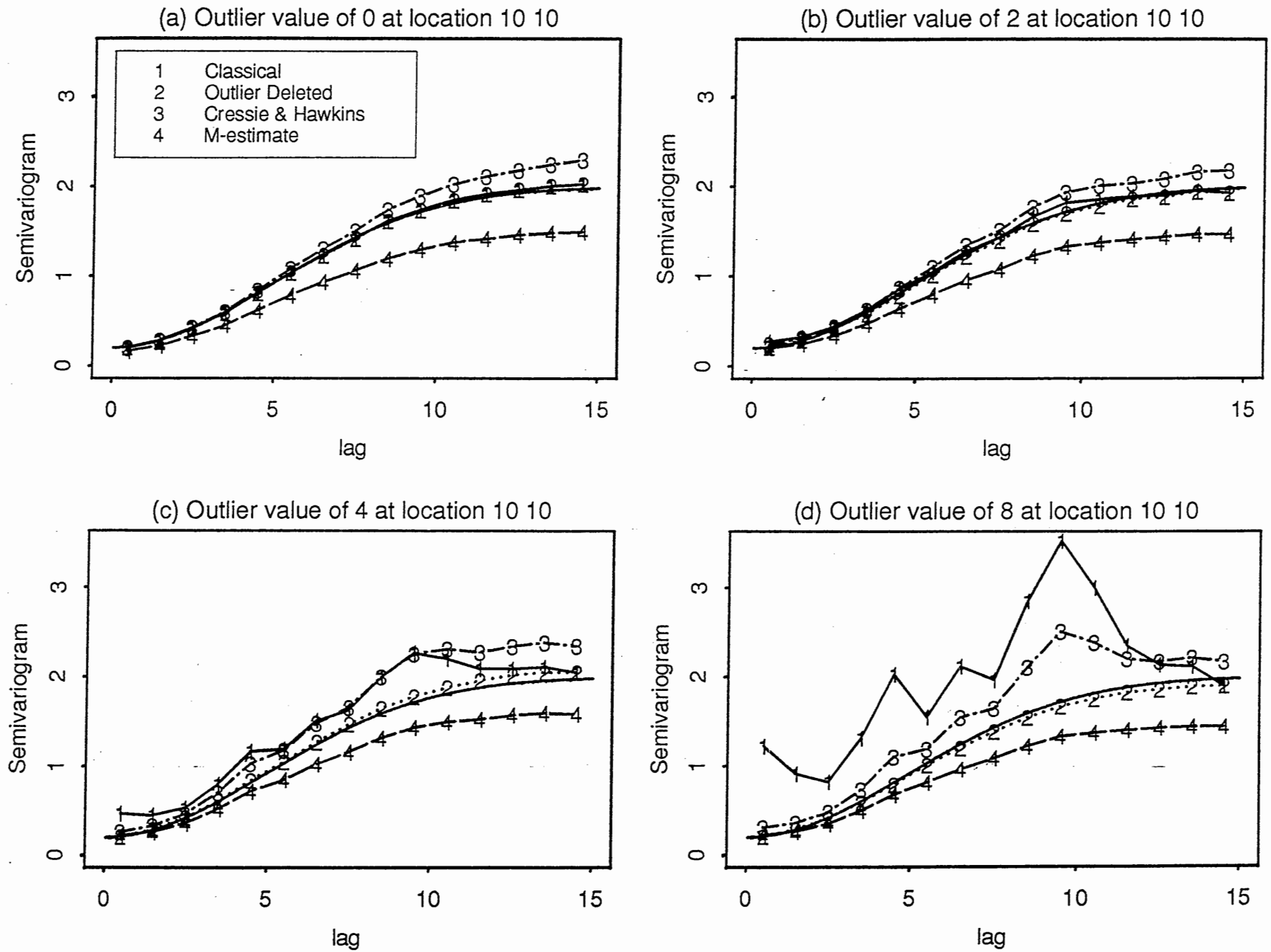
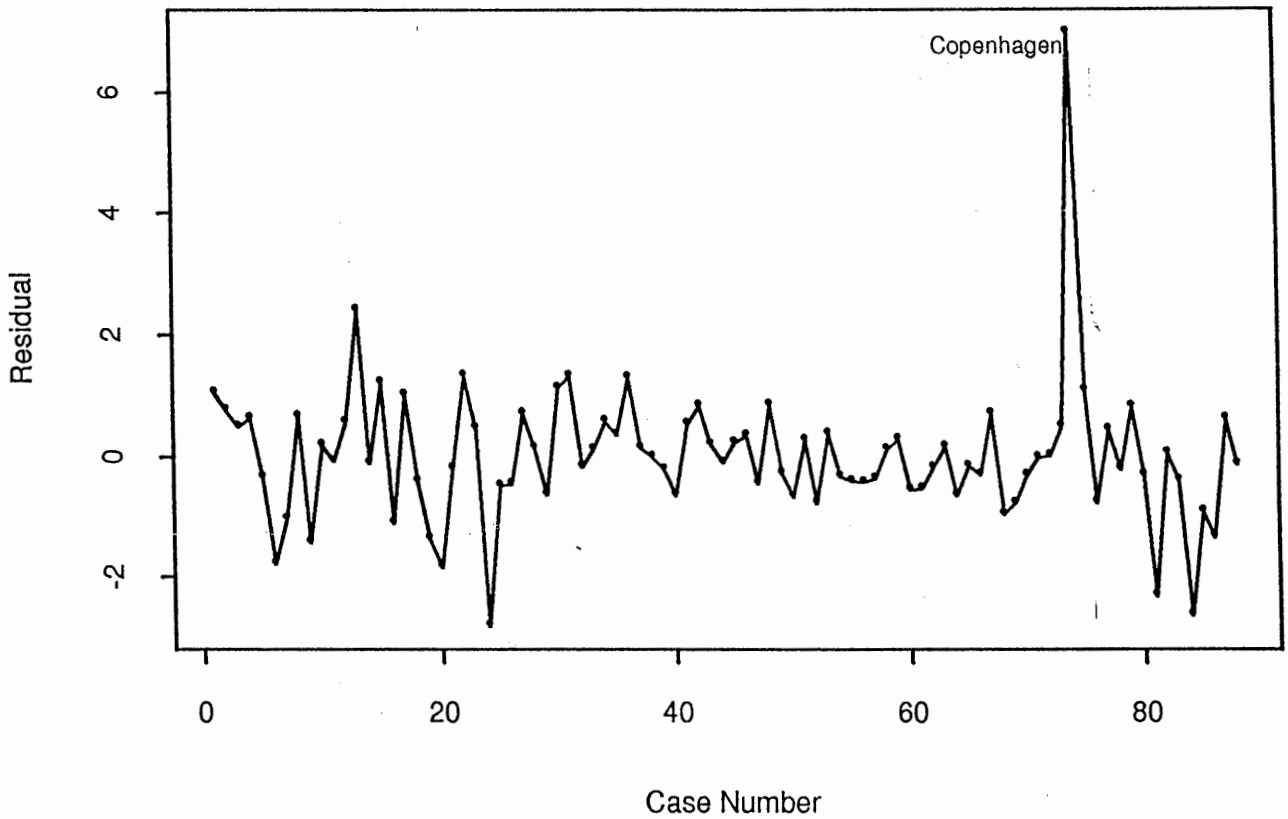


Figure 4. Influence Diagnostics for the November 1990 European Temperature Anomalies.

(a) Studentized Residuals from the Kriging Model Fit.



(b) Model-Parameter Case-Deletion Diagnostics.