Change of support: an inter-disciplinary challenge

C. A. Gotway Crawford[1] and L. J. Young[2]

[1] Centers for Disease Control and Prevention, Atlanta, GA USA
[2] Department of Statistics, University of Florida, Gainesville, FL USA

1 An introduction to change of support in geostatistics

One of the fundamental ideas underlying the field of geostatistics is the concept of a *regularized* variable, the average value of Z(s) over a volume B

$$Z(B) = \frac{1}{|B|}\int_B Z(s)ds, \tag{1}$$

where $|B| = \int_B ds$ is the called the *support* of Z(B). The term support reflects the geometrical size, shape, and spatial orientation of the units or regions associated with the measurements (see e.g., Olea 1991). Changing the support of a variable (typically by averaging or aggregation) creates a new variable. This new variable is related to the original one, but has different statistical and spatial properties. Determining how these properties vary with support is called the *change of support problem*. From the beginning, the field of geostatistics has incorporated solutions to change of support problems (Matheron 1963).

The practical problems driving the initial development of geostatistics were those encountered in the mining industry, with a primary problem being the prediction of the average grade of a mining block from drill core samples. Thus, most change of support problems were concerned with "upscaling," the prediction of a variable whose support is larger than that of the observed data. A common example of this is block kriging where Z(B) is predicted from data $Z(s_1),\ldots, Z(s_n)$ that have mean $E[Z(s)]=\mu$ and semivariogram $\gamma(s-u)=1/2Var[Z(s)-Z(u)]$. The block kriging predictor is given by $\hat{Z}(B) = \sum_{i=1}^{n} \lambda_i Z(s_i)$, where the weights $\{\lambda_i\}$ are obtained by solving the equations (Journel and Huijbregts 1978, Chilès and Delfiner 1999)

$$\sum_{k=1}^{n} \lambda_k \gamma(s_i - s_k) + m = \gamma(B, s_i), \quad i = 1,\ldots,n$$

$$\sum_{k=1}^{n} \lambda_k = 1.$$

Here $\gamma(B, s_i) = \frac{1}{|B|} \int_B \gamma(s_i - u) du$ and m is a Lagrange multiplier from the constrained minimization. The kriging variance is

$$\sigma_K^2(B) = \sum_{i=1}^{n} \lambda_i \gamma(B, s_i) - \gamma(B, B) + m,$$

where

$$\gamma(B, B) = \frac{1}{|B|^2} \int_B \int_B \gamma(s - u) ds du.$$

There are many more sophisticated geostatistical solutions to this change of support problem, including nonlinear methods and those developed to infer the entire probability distribution of the regularized variable (see, e.g., Journel and Huijbregts 1978, Matheron 1984a and b, Cressie 1993b, Rivoirard 1994, Goovaerts 1997, and the compilations in Chilès and Delfiner 1999 and Gotway and Young 2002). However, most practical applications that use them have data of point support (or data measured on small cores or boreholes), and the inferential goal is *upscaling* by *regularization,* so that the inferential goal is prediction of Z(B) (or some function of it) in Eq. 1. Moreover, the volumes B of interest are rectangular blocks and so the integrations required can be done fairly easily and quickly. However, spatial data come in many forms. Instead of measurements associated with point locations, we could have measurements associated with lines, areal regions, surfaces, or volumes. In many disciplines such as geology and soil science, observations often pertain to rock bodies, stratigraphic units, soil maps, and large-scale land use. In geographic and public health studies, the data are often counts or rates obtained as aggregate measures over geopolitical regions such as census enumeration districts and postal code zones. Moreover, the inferential goal may also not be limited to upscaling. For example, modeling hydrological and soil processes often involves making predictions from models that have relatively coarse spatial resolution and these then need to be downscaled to the watershed level or combined with digital elevation data of point support. In many studies in public health, sociology, and political science, the data are counts or rates aggregated over areal regions (e.g., per postal code or per census unit), but individual-level inference is desired. Finally, the idea of regularization as defined through Eq. 1 is not always an appropriate mathematical description of either the data that are available or the inferential quantity of interest. For example, in geographical studies, the data are totals (e.g., the number of people per enumeration district) or rates that are based on population totals and not on the area of the regions. Developing valid inferential methods for upscaling, downscaling and "side-scaling" (in the case of overlapping spatial units) variables is of critical importance to numerous scientific disciplines. It seems natural to try to extend the relatively rich ideology on change of support developed in geostatistics to more general change of support problems.

In this context, we examine the change of support problem from an interdisciplinary point of view. This viewpoint allows us to extract some key ideas, common themes, and general statistical issues common to change of support problems. We provide a brief summary of the various types of solutions that have been proposed to various change of support problems over more than five decades of research conducted in numerous fields of study. The goal of this extroverted contemplation is the search for a general framework for statistical solutions to change of support problems.

2 Why is support important?

Changing the support of a variable through regularization creates a new variable with different statistical and spatial properties. In particular, the variability in Z(B) decreases as the support B increases, the histogram of $Z(B_1)$, …,$Z(B_m)$, $m<n$ will tend to be more symmetric and approximately bell-shaped, and the spatial autocorrelation in the regularized values is altered as well (Journal and Huijbregts 1978, Armstrong 1999). Thus, any inferential procedure must take these factors into consideration. There are numerous examples of this *support effect* in the geostatistical literature, and many methods have been suggested for adjusting for support effects in spatial prediction and resource estimation.

While this view of support has served the mining industry quite well, the situation is more complex in other disciplines. Global Positioning Systems, remote sensing technology, and Geographic Information Systems (GIS) allow greater access to a variety of spatial data and easily permit analysis on almost a limitless choice of spatial units: points, postal code polygons, Census tracts and enumeration districts, hydrogeologic regions, raster images with different pixel sizes, and even regions defined by the whim of the user. More often than not, the data of interest in any one analysis are of different supports that are irregularly shaped. Another factor, related to support, comes into play here: the concept of *scale*. From our review work in this area, we have found that the term is used differently in different disciplines. In fact, few good definitions exist. For example, Bierkens *et al.* (2000) use the terms scale and support interchangeably, defining scale to be support. We argue that they while these two concepts are very much related, they are in fact quite different. From our perspective, spatial scale is defined by both the number and the relative size of the spatial units used to partition a spatial domain of interest. Corresponding to every spatial scale is a level of spatial aggregation that represents the particular mixture of sub-units that comprise the larger units of interest. For a fixed domain, increasing the scale results in a fewer number of larger units. Since size is one aspect of support, clearly support and scale are related. However, we prefer the more general definition of support that includes the shape and orientation and the units. It is possible to partition two spatial domains into subunits with the subunits being of essentially the same size in both partitions, but of different shapes and/or different orientations (Fig. 1).

Geographers have long encountered the interplay between support and scale, noting that the choice of spatial units for analysis is “modifiable,” and that statistical results depend heavily on the way the spatial units are created. In geography, the change of support problem is known as the *Modifiable Areal Unit Problem* (MAUP) (Openshaw and Taylor 1979).

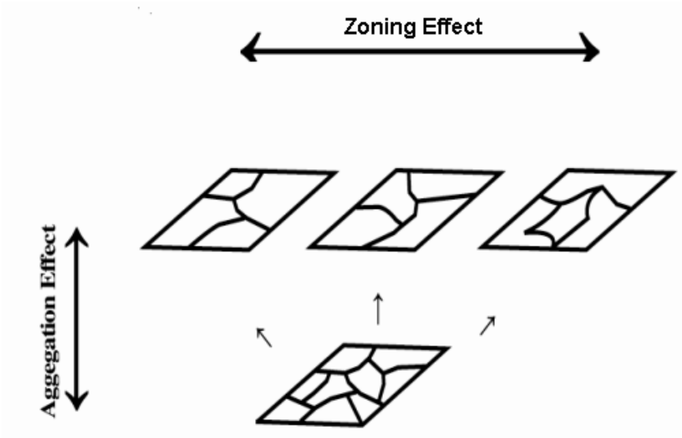

Fig. 1. Components of the support effect and sources of the MAUP. Adapted from Wong(1996).

Thus, the change of support problem and the MAUP are really comprised of two interrelated problems. The first occurs when different inferences are obtained when the same set of data is grouped into increasingly larger areal units. This is often referred to as the *scale effect* or *aggregation effect.* Aggregation reduces heterogeneity among units. The uniqueness of each unit and the dissimilarity among units are both reduced. However, spatial autocorrelation is a mitigating factor: When areal units are similar to begin with, aggregation results in much less information loss than when aggregating highly dissimilar units. Spatial aggregation also affects the spatial variability in the resulting units, often inducing positive spatial autocorrelation, particularly if the aggregation process allows overlapping units. The second, often termed the *grouping effect* or the *zoning effect*, arises from the variability in results due to alternative formations of the areal units that produce units of different shape or orientation at the same or similar scales (Openshaw and Taylor 1979, Wong 1996). The zoning effect is much less pronounced when aggregation of areal units is performed in a non-contiguous or spatially random fashion. It is most apparent only when contiguous units are combined, altering the spatial autocorrelation among the units. Combining smaller units through regularization is analogous to smoothing with different combinations of spatial neighbors. Depending on the similarity of the neighbors, different zoning rules may lead to different analytical results.

In geostatistics, the aggregation effect and the zoning effect are usually treated in a combined fashion through the ideas of the dispersion variance, the regularized semivariogram and its theoretical relationship to the point semivariogram and change of support models that account for both issues simultaneously. However, to appreciate the solutions to the MAUP and the change of support problem developed in other disciplines, we found it helpful to separate the two components. Most solutions to upscaling problems address the effects of aggregation, and most solutions to downscaling problems recognize the need to reconstruct variation at the smaller scale, but the zoning effect issues associated with both of these problems are often ignored.

3 Solutions to change of support problems

Most solutions to change of support problems require spatial prediction of data associated with one set of units based on data associated with another set of units. In developing solutions to change of support problems, the criteria that such predictions should satisfy varies widely across the different disciplines. A collective list of some of the important considerations includes the following:

1. The ability to explicitly account for the differing supports of the spatial units involved;
2. A general framework that can be used for upscaling (aggregation), downscaling (disaggregation), or side-scaling (overlapping units); The framework should allow for upscaling from points to volumes or from volumes to other volumes with larger support. It should allow for downscaling from volumes to volumes with smaller support, or from volumes to points.
3. The predicted surface generated should be smooth across unit boundaries;
4. Standard errors of the predictions can be computed and these should accurately account for the uncertainty involved;
5. Covariates can be used to improve predictions;
6. The method can be used when the data and predictions are averages (as in Eq. 1) or counts/totals;
7. Predictions should lie in the parameter space (e.g., when predicting an inherently positive quantity, the predictions should not be negative);
8. There should be consistency in predictions across scales: For example, consider predicting $Z(A_{ij})$ from data $Z(B_1)$, …,$Z(B_m)$, where the A_{ij} , j=1, … n_i are nested within volume B_i where $A_j \cap A_k=\varnothing$ for $j \neq k$, and $\cup_{j=1}^{n_i} A_j = B_i$. Then the predictions within each volume B should add to the observed datum
$$Z(B_i) = \frac{1}{|B_i|}\sum_{j=1}^{n_i} \hat{Z}(A_{ij})$$
Huang *et al.* (2002) call this the *mass balance* property. When downscaling observed data that are totals and not averages to point support, then the predictions $\hat{Z}(s)$ should satisfy the *pycnophylactic* (volume preserving) property (Tobler 1979):

$$Z(A) = \int_A \hat{Z}(s)ds \cdot$$

9. Ideally, the prediction method should be based on a paucity of model and distributional assumptions;

10. The prediction method should be computationally feasible for routine use within a GIS where it is relatively easily to perform computations involving point-in-polygon operations and digital boundaries.

Of course, asking for a solution that satisfies all of these properties is probably unrealistic. However, this list provides a backdrop against which we can evaluate current solutions and understand their advantages and disadvantages. In the following sections, we provide an overview of some of the general types of solutions to change of support problems and briefly outline some of their main advantages and disadvantages. More comprehensive descriptions of the methods are found in the references provided and many of these are reviewed in more detail in Gotway and Young (2002). We deliberately exclude the rich literature on upscaling and downscaling in many of the physical sciences such as hydrology, soil science, and petroleum engineering in which models that adhere to engineering laws often form a basis for solutions to change of support problems.

3.1 GIS operations and raster calculations

Description: Basic geoprocessing operations with a GIS include union, intersection, and dissolve operations applied to the boundaries of the spatial units in order to create new spatial units. Raster calculations include averaging of interpolated values over irregularly shaped regions ("zonal" statistics) and pixel-by-pixel computations.
Main Advantages: Working with digital boundary files is the consummate utility of GIS. The computations are fast, invisible to the user and can explicitly factor in the support of the different units involved. Layers representing different variables can be combined using raster calculations so that covariates can be incorporated, although the effect of the covariate layers on the predictions must be specified, rather than inferred statistically. Smooth surface generation is straightforward and visualization is automatic.
Main Disadvantages: The main disadvantage is the lack of uncertainty measures for the resulting predictions. Moreover, when several layers with different supports are rasterized to the same cell size and then used in subsequent computations, error propagation is a big concern. Volume-volume disaggregation is done using proportional allocation. Depending on how many operations are used and their nature, the resulting predictions may not be aggregation consistent.

3.2 Spatial smoothing

Description: The goal with spatial smoothing methods is to make a smooth map from aggregated data. Methods in this group vary greatly and include point kriging based on centroids, kernel smoothing (Bracken and Martin 1989), support-adjusted locally weighted regression (Brillinger 1990, Muller *et al.* 1997), and pycnophylactic interpolation (Tobler 1979).
Main Advantages: Point kriging and kernel smoothing based on centroids are easily implemented and provide a measure of uncertainty associated with predictions. The kernel smoothing approach developed by Bracken and Martin (1989) and the pycnophylactic interpolation method of Tobler (1979) computationally constrain the predictions to be aggregation consistent. The methods developed by Brillinger (1990) and Müller *et al.* (1997) are more statistically sophisticated and allow adjustment for covariates and provide a measure of uncertainty. The methods developed by Tobler (1979), Brillinger (1990) and Müller *et al.* (1997) explicitly consider the supports of the units involved.
Main Disadvantages: The major disadvantage to these methods is that are concerned only with the volume-point change of support problem. Constraining predictions to ensure aggregation consistency (as in the methods of Bracken and Martin 1989 and Tobler 1979) makes it difficult to adjust for covariates and to obtain a valid measure of uncertainty. On the other hand, the methods developed by Brillinger (1990) and Müller *et al.* (1997) may not give predictions that are aggregation consistent.

3.3 Regression methods

Description: Proposed by Flowerdew and Green (1992), a regression model is assumed for data associated with "target" units, with the response data on target units treated as missing values. Starting values from proportional allocation are used to obtain initial estimates of the regression parameters. Updated estimates of target-unit data are then obtained from the regression model and constrained to satisfy the pycnophylactic property. This process is repeated until the estimates of the target unit data remain essentially unchanged.
Main Advantages: The main advantage is the ability to use covariates to estimate data on the target units. The regression framework can be used for a variety of change of support problems involving different types of data (binary, discrete and continuous) The computations are fairly simple and could be easily programmed into a GIS script.
Main Disadvantages: Because of the iterative process that includes the pycnophylactic constraint, accurate measures of the uncertainty in target-unit predictions cannot be obtained. Also, the regression model must be built on units formed by the intersection of the target units and the "source" units (those for which data are observed), and so covariates on these "atomic" units must be derived. The support of the units is not considered and spatial autocorrelation is ignored.

3.4 Bayesian hierarchical models

Description: A statistical model is specified for the data, given unknown variables, and then prior distributions are specified for the unknown variables. The unknown variables may include unknown data to be predicted. A posterior distribution is derived from the likelihood of the data that is updated by prior information in accordance with Bayes' theorem. Simulation methods are used to generate realizations from posterior distribution (see, Mugglin and Carlin 1998, Wikle *et al.* 2001, Gelfand *et al.* 2001, Kelsall and Wakefield 2002).
Main Advantages: The methodology is based on very elegant statistical theory combing Bayes' theorem, likelihood estimation and Markov chain theory. The posterior predictive distribution provides a comprehensive description of uncertainty. Complex models that include covariates on many different scales can be more easily constructed hierarchically than simultaneously.
Main Disadvantages: The models are computationally intensive. With the exception of the model in Gelfand *et al.* (2001) each model can be used to solve only one type of change of support problem, and solutions to other problems require complex statistical derivations. Most rely too heavily Gaussian distributions and many account for support only through areal weighting and hence ignore the zonal effect completely. The hierarchical specification can induce unknown constraints within the overall model. There has been little evaluation of the resulting uncertainty distribution (e.g., to assess ergodic properties, or the ability to contain a value of a transfer function of interest as described in Deutsch and Journel 1992 and Gotway and Rutherford 1994).

3.5 Multi-scale tree models

Description: Chou *et al.* (1994) developed a scale-recursive algorithm based on a multilevel tree structure for image processing in engineering. Each level of the tree corresponds to a different spatial scale (see Fig. 2). Data are observed at some of the nodes of the tree and the goal is prediction at other nodes of the tree. Algorithms are based on the Kalman filter. To eliminate some of the artifacts imposed by the tree structure and to ensure mass balance, Huang and Cressie (2000) and Huang *et al.* (2002) extend these models to more general graphical Markov models.
Main Advantages: The recursive nature of the Kalman filter (for which kriging is a special case) is extremely computationally efficient for processing huge data sets. It also provides a measure of uncertainty associated with the predictions.
Main Disadvantages: The tree structure ignores spatial support and it is not clear how it can be adapted to more general cases with overlapping spatial units. Statistical parameter estimation can be difficult.

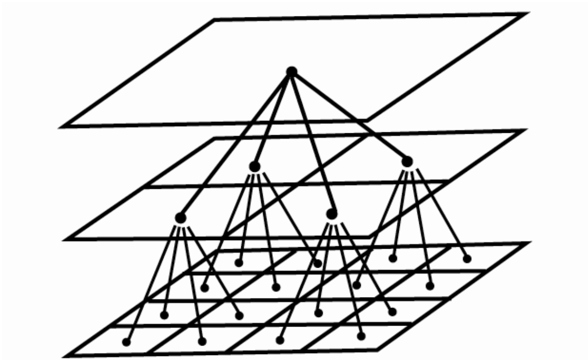

Fig. 2. A tree structure for multiscale processes.

3.6 Geostatistical methods

Description: Includes ``block" kriging, nonlinear geostatistical methods and iso-factorial models (Journel and Huijbregts 1978, Matheron 1984a and b, Cressie 1993b, Rivoirard 1994, Goovaerts 1997, and Chilès and Delfiner 1999).
Main Advantages: The field of geostatistics includes many innovative solutions to change of support problems. These solutions have proven themselves in practical applications such as mining where profitability is of primary concern. A measure of prediction uncertainty can be easily obtained. The basic calculations needed for change of support predictions based on kriging and cokriging can be done in GIS.
Main Disadvantages: Most practical solutions were developed only for the up-scaling problem. Estimating the semivariogram from data that are not of point support may be problematic. Prediction uncertainty may not adequately reflect estimation error in the semivariogram.

4 Towards a general framework

Clearly, the solutions to change of support problems range from those that are simple and require few assumptions, but are statistically unsophisticated (GIS and proportional allocation), to those that are complex and statistically elegant, but require many assumptions and are difficult to implement (Bayesian hierarchical models). Moreover, many solutions are particular to the change of support problem they were developed to address. We seek a compromise, one that provides a unified framework for the different types of change of support problems encountered in a variety of disciplines, is based on fewer assumptions, and can be implemented in a geographic information system (GIS) using current GIS technology, but also one that can incorporate covariates and provide standard errors for the resulting predictions.

While block kriging was developed for the upscaling problem, a slight modification allows the same ideas to be adapted to more general change of support problems (Journel and Huijbregts 1978, Gotway and Young 2002, Gotway and Young 2004). Consider the linear predictor

$$\hat{Z}(A_j) = \sum_{i=1}^{n} w_i(A_j) Z(B_i)$$

based on data $Z(B_1)$, …,$Z(B_n)$, where each weight $w_i(A)$ measures the influence of datum $Z(B_i)$ on the prediction of another variable with differing support, $Z(A)$. The theory of best linear unbiased prediction can be applied to determine optimal weights, $w_i(A)$ in a manner analogous to that used in the development of the block kriging predictor. The key to this development is the relationship between the semivariogram of Z(B) and that of the underlying process Z(s) (Journel and Huijbregts 1978, p. 77)

$$2\gamma(B_{i,}B_j) = \frac{2}{|B_i||B_j|}\int_{B_i}\int_{B_j}\gamma(s-u)dsdu - \frac{1}{|B_i||B_i|}\int_{B_i}\int_{B_i}\gamma(s-u)dsdu - \frac{1}{|B_j||B_j|}\int_{B_j}\int_{B_j}\gamma(s-u)dsdu.$$

Given data of point support, $\gamma(s\text{-}u)$ can be estimated and then used to determine the semivariogram of data at any other support, $\gamma(B_i\ B_j)$ and $\gamma(A_i, A_j)$. Although in many applications, data of point support are available, in man others, such data are not available. However, it is possible to still use this relationship. If a parametric model, $\gamma(s\text{-}u;\theta)$, is assumed for point support semivariogram, an estimate of θ can be obtained, and hence $\gamma(s\text{-}u;\theta)$ can be determined, from data of volume support $Z(B_1)$, …,$Z(B_n)$ (Mockus 1998, Gotway and Young, 2004). Computationally, it is easier to use the covariance functions

$$C(B_i, B_j) = Cov(Z(B_i), Z(B_j)) = \frac{1}{|B_i||B_j|}\int_{\mathbf{B}_j}\int_{\mathbf{B}_i} C(u,v;\theta)dudv$$

since only one multidimensional integration is required. Then, if $Y(B_i)= Z(B_i)-\mu$, θ can be estimated by the value that minimizes (Mockus 1998)

$$\sum_i \sum_j \{Y(B_i)Y(B_j) - \frac{1}{|B_i||B_j|}\int_{B_i}\int_{B_j} C(u-v;\theta)dudv)\}^2$$

The optimal weights needed to construct $\hat{Z}(A)$ can then be obtained from

$$\sum_{k=1}^{n} w_k(A)C(B_i, B_k) - m = C(A, B_i), \quad i = 1,\ldots,n$$

$$\sum_{k=1}^{n} w_k(A) = 1,$$

and the minimized prediction mean squared error (kriging variance) is given by

$$\sigma_K^2(A) = C(A,A) - \sum_{i=1}^{n} w_i(A)C(A,B_i) - m.$$

Gotway and Young (2004) extend these ideas to the “external drift” case where $E[Z(S)] = x(s)'\beta$ and develop an iterative generalize least squares approach to estimating the drift parameters and the autocorrelation parameters simultaneously.

If the data are totals instead of averages, so that $Z^*(B) = \int_B Z(s)ds$, this approach can be used with the normalized variables $N(B) = Z^*(B)/|B|$.

Since an optimal predictor is derived in terms of data with general supports A and B, it can be used for upscaling (spatial aggregation), downscaling (spatial disaggregation), or side scaling (overlapping) units, and the spatial units may be of point, areal, or volumetric support. Because the predictor is linear and honors the data, mass balance properties are inherently satisfied.

However, this approach suffers from the same problems encountered in using geostatistical methods with data of point support, namely the variability in the cross-products if not suitably binned and averaged, and the sensitivity of the estimates of θ to a few large cross-product values and choices for the lag spacing. Another disadvantage of the geostatistical framework when applied to count data is that negative predictions can occur; the predictions are not formally constrained be positive.

5 Summary and challenges

In spite of the rather substantial disadvantages associated with using GIS operations to combine spatial data, the ability to easily implement solutions to change of support problems within a GIS is overwhelmingly appealing. Thus, overall, this approach is the most commonly used method for combining incompatible spatial data and solving complex change of support problems. While Bayesian hierarchical models and isofactorial models offer elegant statistical solutions to a variety of change of support problems, their complexity (both statistical and computational) and their dependence on a large number of unverifiable, pedantic assumptions make them unattractive for routine use in most applied sciences at the present time. Thus, as a compromise, we considered a geostatistical approach to general change of support problems that allows downscaling and side scaling. This approach explicitly accounts for the supports of the data, can incorporate covariate information to improve the predictions, and provides a measure of uncertainty for each prediction.

While the geostatistical framework presented here is not without its disadvantages, it offers a way to put the concept of spatial support back into spatial analysis. Subsequent research and development could easily adapt this framework for use as a routine part of many software packages.

References

Armstrong M (1999) Basic Linear Geostatistics. Springer-Verlag New York

Bierkens MFP, Finke PA, De Willigen P (2000) Upscaling and Downscaling Methods for Environmental Research. Kluwer Dordrecht

Bracken I., Martin D (1989) The generation of spatial population distributions from census centroid data. Environment and Planning A 21: 537–543

Brillinger DR (1990) Spatial-temporal modeling of spatially aggregate birth data. Survey Methodology 16: 255-269

Chilès JP, Delfiner P (1999) Geostatistics: Modeling Spatial Uncertainty. John Wiley & Sons, New York

Chou KC, Willsky AS, Nikoukhah R (1994) Multiscale systems, Kalman filters, and Riccati equations. IEEE Transactions on Automatic Control, 39: 479-492

Cressie N (1993) Statistics for Spatial Data. John Wiley & Sons, New York

Cressie N (1993b) Aggregation in geostatistical problems. In Soares, A (ed.) Geostatistics Troia '92, Kluwer Academic Publishers, Dordrecht, 25-35

Deutsch CV, Journel AG (1992) GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York.

Flowerdew R and Green M (1992) Developments in areal interpolating methods and GIS. Annals of Regional Science, 26: 67-78

Gelfand AE, Zhu L, Carlin BP (2001) On the change of support problem for spatio-temporal data. Biostatistics, 2: 31-45

Goovaerts P (1997) Geostatistics for Natural Resources Evaluation. Oxford University Press, New York

Gotway CA and Rutherford BM (1994) Stochastic simulation for imaging spatial uncertainty:comparison and evaluation of available algorithms. In Armstrong, M and Dowd, PA (eds). Geostatistical Simulations, Kluwer: Dordrecht, 1-21

Gotway CA, Young LJ (2002) Combining incompatible spatial data. Journal of the American Statistical Association, 97: 632-648

Gotway CA, Young LJ (2004) A geostatistical approach to linking geographically-aggregated data from different sources. Technical report # 2004-012, Department of Statistics, University of Florida

Huang H-C, Cressie N (2001) Multiscale graphical modeling in space: Applications to command and control. In Moore, M. (ed.) Spatial statistics. Methodological Aspects and Some Applications, vol. 159 of Lecture notes in Statistics, Springer Verlag, New York

Huang H-C, Cressie N, Gabrosek J (2002) Fast resolution-consistent spatial prediction of global processses from satellite data. Journal of Computational and Graphical Statistics, 11:1-26

Journel AG, Huijbregts CJ (1978) Mining Geostatistics. Academic Press, London

Kelsall J, Wakefield J (2002) Modeling spatial variation in disease risk: A geostatistical approach. Journal of the American Statistical Association, 97: 692-701

Matheron G (1963) Principles of geostatistics. Economic Geology, 58: 1246-1266

Matheron G (1984a) Isofactorial models and change of support. In Verly, G. *et al.* (eds.) Geostatistics for Natural Resources Characterization, Reidel, Dordrecht, 449-467

Matheron G (1984b) The selectivity of the distributions and the "second principle of geostatistics." In Verly, G. *et al.* (eds.) Geostatistics for Natural Resources Characterization, Reidel, Dordrecht, 449-467

Mockus A (1998) Estimating dependencies from spatial averages. Journal of Computational and Graphical Statistics, 7: 501-513

Mugglin AS, Carlin BP (1998) Hierarchical modeling in geographic information systems: Population interpolation over incompatible zones. Journal of Agricultural, Biological, and Environmental Statistics, 3: 117-130

Müller HG, Stadtmüller U, Tabnak F (1997) Spatial smoothing of geographically aggregated data, with application to the construction of incidence maps. Journal of the American Statistical Association, 92: 61-71

Olea RA (ed.) (1991) Geostatistical Glossary and Multilingual Dictionary. Oxford University Press, New York

Openshaw S, Taylor P (1979) A million or so correlation coefficients. In Wrigley, N. (ed.) Statistical Methods in the Spatial Sciences, Pion, London, 127-144

Rivoirard J (1994) Introduction to disjunctive kriging and non-linear geostatistics. Clarendon Press, Oxford

Tobler W (1979) Smooth pycnophylactic interpolation for geographical regions (with discussion). Journal of the American Statistical Association, 74: 519-536

Wikle CK, Milliff RF, Nychka D, Berliner LM (2001) Spatio-temporal hierarchical Bayesian modeling: Tropical ocean surface winds. Journal of American Statistical Association, 96: 382-397

Wong DWS (1996) Aggregation effects in geo-referenced data. In Griffiths, D (ed.) Advanced Spatial Statistics, CRC Press, Boca Raton, Florida, 83-106