Title
Incorporating Model Uncertainty into Spatial Predictions

ISBN
9781461384946

Author
Handcock, Mark S

Publication Date
1996

DOI
10.1007/978-1-4613-8492-2_8

Peer reviewed

In "Environmental Studies: Mathematical, Computational and Statistical Analysis".
IMA Volumes in Mathematics and its Applications, Vol. 79, 189-200.
Edited by Mary F. Wheeler. Springer-Verlag, New York.

INCORPORATING MODEL UNCERTAINTY INTO SPATIAL PREDICTIONS

MARK S. HANDCOCK DEPARTMENT OF STATISTICS & OPERATIONS RESEARCH, NEW YORK UNIVERSITY, LEONARD N. STERN SCHOOL OF BUSINESS, NEW YORK, NY 10012. *

Abstract. We consider a modeling approach for spatially distributed data. We are concerned with aspects of statistical inference for Gaussian random fields when the ultimate objective is to predict the value of the random field at unobserved locations. However the exact statistical model is seldom known before hand and is usually estimated from the very same data relative to which the predictions are made. Our objective is to assess the effect of the fact that the model is estimated, rather than known, on the prediction and the associated prediction uncertainty. We describe a method for achieving this objective. We, in essence, consider the best linear unbiased prediction procedure based on the model within a Bayesian framework.

These ideas are implemented for the spring temperature over the region in the northern United States based on the stations in the United States historical climatological network reported in Karl, Williams, Quinlan & Boden (1990).

Key words. Gaussian random fields; Bayesian statistics; Climatic change.

1. Introduction. We develop spatial-temporal models for the spring temperature over a region in the northern United States covering eastern Montana through the Dakotas ($90° - 107°$ in longitude) and northern Nebraska up to the Canadian border ($41° - 49°$ in latitude). The empirical work of Lettenmaier, Wood and Wallis (1992) suggests that the temperatures over the spring period might exhibit a temporal pattern not found in the winter months. In addition the relatively stable and simple topography of the region help to ensure homogeneity and the minimization of localized effects.

The traditional best linear unbiased prediction procedure ("Kriging") is used in this paper for inference, but within a Bayesian framework. Particular attention is paid to the treatment of parameters in the covariance structure and their effect on the quality of the prediction. Our approach is to exam how posterior predictive distributions of areal quantities change over time. The objective is to see if there have been changes in areal temperature that are discernible from the year-to-year variation.

A companion study reported in Handcock & Wallis (1992) considered the winter months and this region because GCM predictions of climatic change ($4°F - 10°F$) induced by increased greenhouse gases are expected to be at maximum for high latitudes during the winter months (Mitchell (1989), IPCC (1990)). However there was no indication that the areal mean temperature for this time of the year in this region has changed over the last half century. There the posterior predictive distributions were used

*This paper has benefited greatly from the guidance and suggestions of Michael Stein.

as a basis for calibrating temperature shifts by the historical record. In particular, the objective was to understand how soon gradual increases in temperature over this region would be discernible from the year-to-year variation. A similar analysis could be undertaken for the spring season considered here.

There has been much interest recently in climatic change and potential global warming. Of central focus is the phenomenon popularly called the "greenhouse effect": the heating of the earth via the entrapment, by certain gases, of long-wave radiation emitted from the earth's surface. This effect produces a global mean temperature of about $59°F$ rather than an estimated $-6°F$ in the absence of atmosphere (Mitchell (1989)). Increasing concentrations of the gases thought to contribute to this effect have led to concern in the scientific community about increases in temperature and the resulting climatic effects.

There appears to be no clear cut consensus on the extent of global warming over the last century. Most estimates run from $0.5°F$ to $1.0°F$. The difficulty is the lack of good long-term data over large regions. The global temperature constantly changes on time scales of tens of thousands of years. In fact there have been times in the past millennium when it has been much warmer than the majority of global warming scenarios. The question here is over a rapid change over the next century that will have enormous impact on the environment.

Much of the evidence for a global warming effect has been based on large-scale Global Circulation Models (GCMs). These use multi-level mathematical representations of the atmosphere for weather prediction. Given the complexity of the environment and the relative simplicity of the models there is much controversy concerning their validity. Results from the four most widely cited GCMs (1) the National Center for Atmospheric Research (NCAR), (2) Geophysical Fluid Dynamics Laboratory (GFDL) of the National Oceanographic and Atmospheric Administration, (3) the Goddard Institute of Space Studies (GISS) and, (4) the Hadley Center for Climate Prediction and Research at Bracknell, England, are still far from being in agreement, although all models predict higher winter temperatures at the higher northern altitudes as a function of increasing greenhouse gases.

The modeling framework is developed in Section 2. The application of the model to the spring temperature field is considered in Section 3. The areal mean spring temperature is modelled in Section 4.

2. Methodology. Suppose $Z(x)$ is a real-valued stationary Gaussian random field on R with mean

$$E\{Z(x)\} = f'(x)\beta,$$

where $f(x) = \big\{f_1(x), \dots, f_q(x)\big\}'$ is a known vector-valued function and β is a vector of unknown regression coefficients. Furthermore, the covariance function is represented by

$$\mathrm{cov}\big\{Z(x), Z(y)\big\} = \alpha K_\theta(x, y) \qquad \text{for } x, y \in R$$

where $\alpha > 0$ is a scale parameter, $\theta \in \Theta$ is a $p \times 1$ vector of structural parameters and Θ is an open set in $\mathbb{R}^p$. In the general case, we observe $\{Z(x_1), \ldots, Z(x_n)\}' = Z$ and will focus on the prediction of $Z(x_0)$. In our application $x_1, \ldots, x_n$ are the spatial locations of the stations in the network. We will focus on the prediction of $Z(x_0)$, where x_0 is a new location in the region of interest. The Kriging predictor is the best linear unbiased predictor of the form $\widehat{Z}_\theta(x_0) = \lambda'(\theta)Z$; that is, the unbiased linear combination of the observations that minimizes the variance of the prediction error. It is straightforward to show that the corresponding weight vector $\lambda(\theta)$ defining $\widehat{Z}_\theta(x_0)$ is given by

$$\lambda'(\theta) = k_\theta' K_\theta^{-1} + b_\theta'(F'K_\theta^{-1}F)^{-1}F'K_\theta^{-1}, \tag{1}$$

where

$$\begin{aligned} F &= \{f_j(x_i)\}_{n\times q}, \\ k_\theta &= \{K_\theta(x_0, x_i)\}_{n\times 1}, \\ K_\theta &= \{K_\theta(x_i, x_j)\}_{n\times n}, \\ b_\theta &= f(x_0) - F'K_\theta^{-1}k_\theta. \end{aligned}$$

In the example $x = (x_1,\ x_2)$ and we can take $f_1(x) = x_1$ and $f_2(x) = x_2$, the latitude and longitude of locations within the geographic region, respectively. A third component of the mean will be added in Section 3. The covariance function represents the covariance between the temperature at the locations $x = (x_1,\ x_2)$ and $y = (y_1,\ y_2)$.

Note that the prediction weights $\lambda'(\theta)$ do not depend on α or β. Under our Gaussian model, for fixed α, β, and θ, the conditional distribution of $Z(x_0)$ is

$$Z(x_0) \mid Z \sim N\Big(\ k_\theta' K_\theta^{-1} Z + b_\theta' \beta,$$

where $\hat{\beta}(\theta) = (F'K_\theta^{-1}F)^{-1}F'K_\theta^{-1}Z$ and $N(\cdot,\cdot)$ denotes the Gaussian distribution. The prediction error, $Z(x_0) - Z(x_0)$, then is mean-zero Gaussian with variance $V_\theta = K_\theta(x_0, x_0) - k_\theta' K_\theta^{-1} k_\theta + b_\theta'(F'K_\theta^{-1}F)^{-1}b_\theta$ and αV_θ as given in Ripley (1981).

In traditional Kriging, one estimates α, β and θ by either likelihood methods or various *ad hoc* approaches. The likelihood approach to the estimation of the covariance structure was first applied in the hydrological and geological fields following Kitanidis (1983), Kitanidis and Lane (1985) and Hoeksema and Kitanidis (1985). Mardia and Marshall (1984) is a standard reference in the statistical literature. As described below, the behavior of the predictor is usually estimated by "plugging-in" the estimates into 1 and the prediction error variance. If θ is known so that only the location parameter β and the scale parameter α are uncertain then we are in a standard generalized least-squares setting. The distinction between the generalized least squares setting and the random field setting is the uncertainty in the structural parameter θ.

As β is a location parameter we expect that our prior opinions about β bear no relationship to those about α and *a priori* might expect α and β to be independent, leading to the use of Jeffreys's prior. Partly for convenience, the form of the prior used here will be

$$\mathrm{pr}(\alpha,\ \beta,\theta) \propto \mathrm{pr}(\theta)/\alpha$$

It easily follows from Zellner (1971) that the predictive distribution of $Z(x_0)$ conditional on θ and Z is

$$Z(x_0) \mid \theta, Z \sim t_{n-q}\left(\widehat{Z_\theta}(x_0),\quad \frac{n}{n-q}\widehat{\alpha}(\theta)V_\theta \right), \tag{2}$$

a shifted t distribution on $n-q$ degrees of freedom.

The marginal posterior distribution of θ can be shown to be

$$\mathrm{pr}(\theta \mid Z) \propto \mathrm{pr}(\theta)\cdot |K_\theta|^{-1/2}|F'K_\theta^{-1}F|^{-1/2}\widehat{\alpha}(\theta)^{-(n-q)/2} \tag{3}$$

The Bayesian predictive distribution for $Z(x_0)$ is

$$\mathrm{pr}(Z(x_0) \mid Z) \propto \int_\Theta \mathrm{pr}(Z(x_0) \mid \theta, Z)\cdot \mathrm{pr}(\theta \mid Z) d\theta$$

where the integrand is given by 2 and 3 As the dependence of K_θ on θ is not specified this expression can not be simplified and further exploration will in general require numerical computation. If prior information is available it may be directly incorporated into 3, although additional numerical integration may be necessary if prior dependencies among (α,β,θ) are envisaged.

Suppose we use an estimation procedure to select the parameters $(\tilde{\alpha},\tilde{\theta})$ of a covariance structure. These may be arrived at by any procedure, although the usual methods are maximum likelihood, weighted least squares or derived from empirical correlation functions. The distribution that an investigator would use as a basis for inference about $Z(x_0)$ would be

$$Z(x_0) \mid \tilde{\alpha},\tilde{\theta}, Z \sim N\Big(\widehat{Z_{\tilde{\theta}}}(x_0),\ \tilde{\alpha}V_{\tilde{\theta}} \Big),$$

plugging in $(\tilde{\alpha},\tilde{\theta})$ for (α,θ) in the conditional distribution of $Z(x_0)$.

The difficulty of using a point estimate of the covariance structure as a surrogate for the "true" covariance structure is that the uncertainty in the estimate is not directly translated to the final inference. The maximum likelihood estimate may be the best single representative available, but this reduction itself can be detrimental to the inference.

One approach to include the uncertainty in the estimate is to use the Bayesian framework and base inference on the posterior distributions of the quantities of interest. This approach takes into account the complete likelihood surface rather than only plugging in the maximum likelihood estimate of the covariance structure. It allows the performance of the usual plug-in predictive distribution based on an estimated covariance structure to be critiqued within a larger framework.

Depending on the influence of θ on the spread and location of $\mathrm{pr}(Z(x_0) \mid \theta, Z)$, the Bayesian predictive distribution might be wider or narrower than the plug-in predictive distribution. The location of the plug-in predictive distribution may also be quite different from the Bayesian predictive distribution. Typically the Bayesian predictive distribution will have no simple analytic form and must be determined numerically. The difference between the plug-in and Bayesian predictive distributions represents the difference in inference between the traditional Kriging approach and the full Bayesian approach.

3. Modeling spring temperature fields. In this section we consider a spatial model appropriate for a meteorological field over a single time period. The field discussed here is the average spring temperature. The daily average temperature at a location is defined to be the mean of the daily maximum and the daily minimum at that location. The average spring temperature is defined to be the average daily average temperature over the months March, April and May.

The basis of the data is a network of 1219 stations (the HCN network) for the contiguous United States developed by the U. S. Carbon Dioxide Information Analysis Center "with the objective of compiling a data-set suitable for the detection of climatic change" (Karl, Williams, Quinlan & Boden (1990)). The actual data values are from Wallis, Lettenmaier and Wood (1991) which applied a enhanced method of adjusting for missing days (See Handcock & Wallis (1992)).

The components of the parametric mean function, $f_i(\cdot)$, should clearly include the latitude, longitude and elevation of each station. Other possibilities are polynomials in latitude, longitude, elevation, and the distance to the closest urban area or transformations of them. The ultimate choices for components for the mean function were latitude, longitude and elevation, as additional components did not have an appreciable effect on the likelihood ratios or on the likelihood function itself.

The parametric family of covariance functions used in this analysis is the Matérn class discussed in Matérn (1986) and Handcock & Stein (1992) that we feel provides a sound foundation for the parametric modeling of Gaussian random fields. In the form used here it is spatially isotropic and homogeneous: $K_\theta(x,y) \equiv K_\theta(|x-y|)$ is usually expressed as a function of a single scalar variable:

$$\frac{1}{2^{\theta_2-1}\Gamma(\theta_2)} \cdot \left(\frac{x}{\theta_1'}\right)^{\theta_2} \mathcal{K}_{\theta_2}\left(\frac{x}{\theta_1'}\right)$$

where $\theta_1' = \theta_1/(2\sqrt{\theta_2})$ and $\theta_1 > 0$ is a scale parameter controlling the range of correlation. The smoothness of the field is controlled by $\theta_2 > 0$. $\mathcal{K}_{\theta_2}$ is the modified Bessel function of order θ_2 discussed in Abramowitz and Stegun (1964), §9.

The class is motivated by the smooth nature of the spectral density, the wide range of behaviors covered and the interpretability of the parameters. The Exponential class corresponds to the sub-class with smoothness parameter $\theta_2 = 1/2$, that is

$$K_E(x) = \theta_1 \exp(-x/\theta_1).$$

As $\theta_2 \to \infty,\ K_\theta(x) \to \exp(-x^2/\theta_1^2)$, often called the "Gaussian" covariance function. We shall refer to it as the Squared Exponential model. This model forms the upper limit of smoothness in the class. A general treatment is given in the seminal work by Matérn (1986).

Based on meteorological arguments, we believe that the underlying meteorological field is continuous and may be differentiable many times . The magnitude of both random and systematic measurement error ("nugget effect") varies from year-to-year. It can be incorporated by adding a single additional parameter (θ_3) to the covariance function:

$$\operatorname{cov}\big\{Z(x), Z(y)\big\} = \alpha\Bigg(\theta_3\mathcal{I}(x=y)\ +\ K_\theta(x,y)\Bigg)$$

where $\mathcal{I}(\cdot)$ is the indicator function. For example, for 1984 the maximum likelihood estimate had no nugget effect ($\widehat{\theta_3} = 0$). Typically the estimated nugget effect was 25% - 50% of the point variance. In contrast, the mean winter temperatures reported in Handcock & Wallis (1992) the nugget effect appeared to be small relative to the year to year variation. As expected the estimate of the smoothness of the field increases when a nugget effect is included. For example, for 1988 the maximum likelihood estimate for the covariance structure based on the Matérn class is $(\widehat{\alpha}, \widehat{\theta}) = (1.33°F^2,\ 1.02°,\ 10.3,\ 77\%)$. The range of dependence (1.02°) spans approximately a ninth of the region under study. The point standard deviation of the mean spring temperature is $\sqrt{\widehat{\alpha}\ +\ \widehat{\theta_3}} = 1.5°F$. The smoothness of the field is estimated to be ten mean-square derivatives. Under the Gaussian assumption this implies that the observed field has ten derivatives. The likelihood is very flat for $\theta_2 > 1$. The maximum likelihood Squared Exponential model ($\theta_2 = \infty$) with a nugget effect is $(\widehat{\alpha}, \widehat{\theta}) = (1.33°F^2,\ 0.97°,\ \infty,\ 76\%)$, with only a 0.01 decrease in log-likelihood from the previous model. This feature was typical of most years: the log-likelihood is very flat for large smoothness values indicating that there is little information in the spatial observations to discriminate between these smoothnesses *in the presence of a substantial micro-scale variation.*

The estimates of the regression parameters indicate that the mean spring temperature decreases about $1.4°F$ per degree increase in latitude. There is a $0.4°F$ decrease per 1000 feet increase in elevation. In addition for every degree increase in longitude eastward the mean spring temperature decreases about $1.0°F$. This last effect is possibly a surrogate for the spring climatic patterns over the region.

The prior used here for θ_1 and θ_2 reasonably assumes prior independence between the smoothness and the range It is marginally non-informative for $\frac{\theta_i}{1+\theta_i}$ on $[0,\ 1]$ $i = 1, 2$. The prior reflects the belief that higher values of the smoothness and range are *a priori* less likely than smaller values. It should be emphasized that the method is designed to incorporate an informative prior if the meteorologist, hydrologist and/or statistician has one.

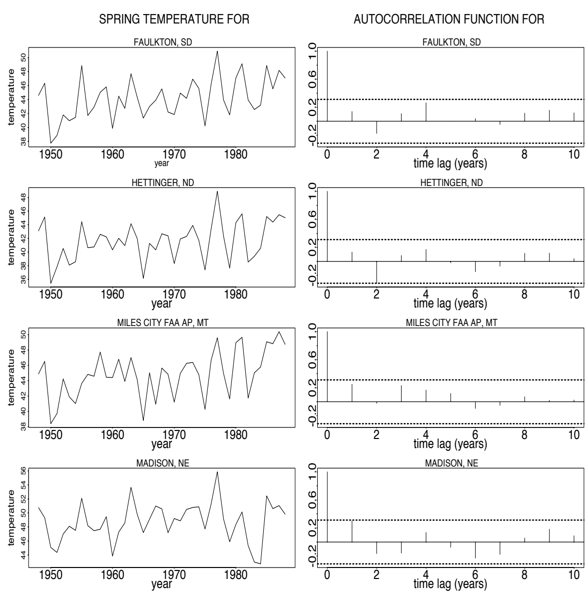

FIG. 1. *Time-series and empirical autocorrelation functions for four typical sites.*

This spatial analysis was repeated independently for each year of data from 1948-88.

We can add the temporal component of the model, generalizing the random field to $Z_t(x)$ where $t = 1948, \ldots$ represents the spring of observation. We consider the time-series of data from each station, independent of the spatial information.

Figure 3.1 presents the time-series and empirical autocorrelation functions for four spatially separate stations. Individually the time-series are quite variable over time. Inspection of the series as a whole suggests a mild upward trend over the last half century time. The right hand side figures are the sample autocorrelation functions corresponding to the time-series. The dashed boundaries represent approximate 95% confidence limits. Note the similar patterns in the series over time and the lack of first lag autocorrelation.

We investigated these fields for short-memory temporal structure. We found little evidence for AR(1) structure and indication that the series are close to uncorrelated over time. We also considered the presence of significant dependence between observations a long time span apart, postulating an autoregressive integrated moving average processes with non-integral degrees of differencing, d. (ARIMA(p, d, q)). There was little (likelihood) evidence for such structure.

4. Measuring areal mean temperature. In the previous sections we found a complex spatial structure to the mean spring temperatures, little temporal dependence structure and little evidence for changing spatial structure over time. In this section we focus interest in a measure of the areal mean temperature over the region of interest. The time-series of areal mean temperatures is defined by:

$$\bar{Z}_t = \frac{1}{|R|} \int_R Z_t(x)dx \qquad t = 1948, \ldots, 1988, \ldots$$

where $|R|$ is the area of the region R. Thus at each point in time, $\bar{Z}_t$ represents the average temperature over the region and is a function of the field $Z(x)$. $\bar{Z}_t$ provides a natural measure for the detection of changing climatic patterns over the region. As the region is devoid of gross topographic features, it provides a convenient measure of overall temperature during the spring. It is important to note that $\bar{Z}_t$ is a characteristic of the temperature field itself, and not a characteristic of the stations in the network. The behavior of the areal mean temperature will provide an indication of the overall changes in climate over the region independent of the individual stations.

Based on our model, we can summarize the available information for $\bar{Z}_t$ from the predictive density $P(\bar{Z}_t \mid Z_{1948},\ Z_{1949}, \ldots, Z_{1988})$, that is, the posterior density of $\bar{Z}_t$ given the complete spatial-temporal information available. The evidence in the previous section indicate that the temporal dependence is weak so that $P(\bar{Z}_t \mid Z_{1948},\ Z_{1949}, \ldots, Z_{1988})$ is very well approximated by $P(\bar{Z}_t \mid Z_t)$. This also has the advantage of relative computational tractability.

Now

$$P(\bar{Z}_t \mid Z_t) \propto \int_\Theta P(\bar{Z}_t \mid \theta, Z_t) \cdot P(\theta \mid Z_t) \;\; d\theta \tag{4}$$

where $P(\bar{Z}_t \mid \theta, Z_t)$ is the predictive distribution conditional on θ and Z_t, and $P(\theta \mid Z_t)$ is the posterior distribution for the structural parameter for the year t given in 3. This calculation requires two-dimensional numerical integration for each year. By plotting the predictive distributions over time we can observe how our knowledge of the areal mean temperature changes based on how the information in the network changes.

How can we further summarize the areal mean temperature? The distributions are symmetric and have a similar t-like distributional shape. The ratio of largest to smallest variance is 2.2. To further explore the temporal changes in $\bar{Z}_t$ we will consider the time-series of maximum *a posteriori* (MAP) values from 4. While this clearly represents a reduction in information relative to the full distribution, it facilitates examination.

4.1. Temporal structure of the mean areal temperature. Figure 4.1 represents the MAP values for the last half century. Note the suggestion of a trend over time. Some interesting years have been indicated.

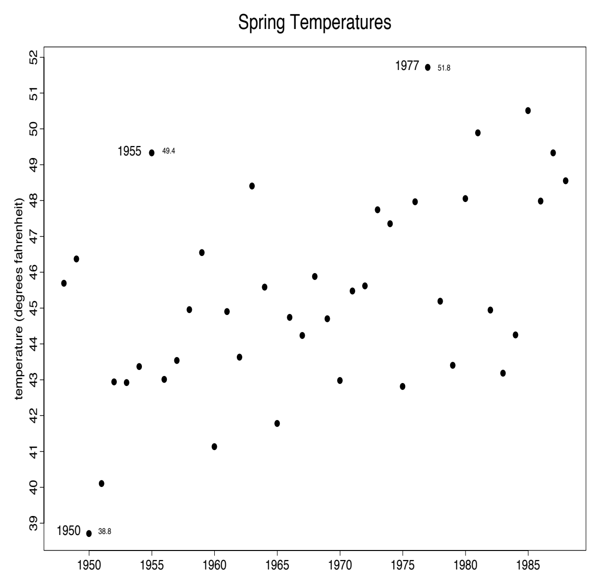

FIG. 2. *This is the time-series of (the MAP estimates for the) areal mean temperature. There is little short or long term dependence.*

Using the same approach as §3 there was little evidence, over this period, of short or long term dependence in the MAP values. However long term dependence in climatological series can occur over time scales of a centuries or more and such dependence would not be apparent from our half century of observation. The trend apparent in Figure 4.1 could simply be an artifact of such long term dependence.

4.2. A static model for mean areal temperature. A reasonable model for the mean areal temperature over the last half century is

$$\bar{Z}_t = \mu_t + \epsilon_t \quad t = 1948, \ldots, 1988, \ldots \tag{5}$$

where $\{\epsilon_t\}_{t=1948}^{1988}$ is an independent, and identically distributed Gaussian sequence with zero mean and variance σ^2. The sequence $\{\mu_t\}_{t=1948}^{1988}$ represents the mean level. The motivation is the absence of strong temporal dependence and the approximate constant variances. One base line model is that the means are temporally stable: $\mu_t \equiv \mu$. Here μ will be called the static areal mean temperature.

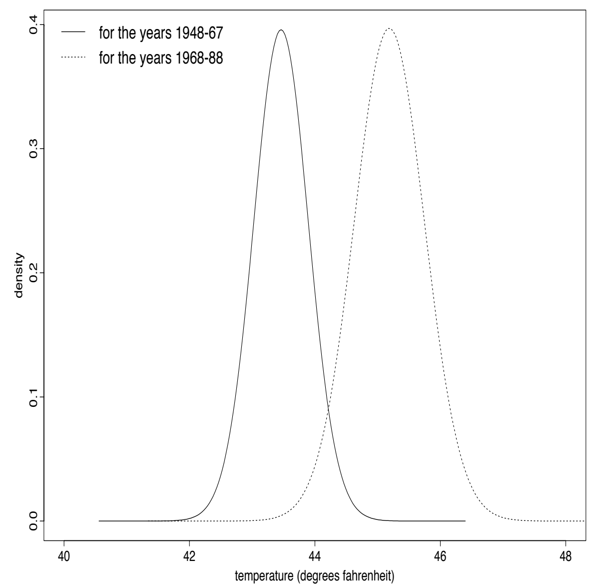

FIG. 3. *Predictive distributions for the static areal mean temperature for two separate time periods. The solid line refers to 1948-67, while the dashed line refers to 1968-88.*

Under the model 5, the posterior distribution for μ over time periods $Z_{t_1}, \ldots, Z_{t_T}$, $P(\mu \mid Z_{t_1}, \ldots, Z_{t_T})$, is

$$\int_{\bar{Z}_{t_1}, \ldots, \bar{Z}_{t_T}} P(\mu \mid \bar{Z}_{t_k}) \cdot P(\bar{Z}_{t_k} \mid Z_{t_k}) \;\; d\bar{Z}_{t_1} \ldots d\bar{Z}_{t_T}$$

The first part of the integrand can be directly calculated and the latter is available from 4. As in the previous case, this requires numerical integration. Figure 4.2 represents the predictive distributions for the static areal mean temperature for two separate time periods. They summarize our uncertainty about the static areal mean temperature over these periods. The solid line refers to 1948-67, while the dashed line refers to 1968-88. The mean for the first period is $43.5°F$ with a standard deviation of $1.0°F$. The mean for the second period is $45.3°F$ with a standard deviation of $1.4°F$. Note that the distributions have a small overlap, with the information from the latter years indicating warmer temperatures. The uncertainty in the mean areal temperature for the latter period is larger, reflecting the increased variation in the annual values.

A simple alternative model is

$$\mu_t \quad = \quad \mu \; + \; \beta(t - 1948) \quad t = 1948, \ldots, 1988, \ldots$$

The posterior distribution of β is centered about $1.2°F$ per decade and the vast majority of its mass is on positive values. However this assumes a linear trend which, as discussed in the previous section, is confounded with any long term dependence over a time scale much longer than the forty year observation period.

5. Conclusion. The Kriging procedure is often described as optimal (Matheron (1971)) because it produces optimal predictions when the covariance structure of the random field is known. If the covariance structure is not known and needs to be estimated, it is then necessary to assess the effect of the estimation on the prediction and the associated prediction uncertainty.

The approach in this paper takes into account the uncertainty about the covariance function expressed in the likelihood surface and ignored by point estimates of the covariance function.

The application to modeling spring temperature indicates that there is significant micro-scale variation over a spatially smooth field. There is substantial variation in the spring temperature from year-to-year that is spatially correlated (Figure 3.1). The areal mean temperature is correspondingly variable and appears to be increasing over the period considered (Figure 4.1). However, this finding has been confirmed using different statistical approaches (Lettenmaier, Wood and Wallis (1992)). It is also evident from this later study that, had other regions or periods been picked, the results would have been quite different. When the model parameter uncertainty is incorporated for the increase in areal mean temperature is still discernible (Figure 4.2).

REFERENCES

Abramowitz, M. & Stegun, I.A. (1965). *Handbook of Mathematical Functions.* New York: Dover.

Handcock, M.S. and Stein, M.L. (1993). *A Bayesian Analysis of Kriging.* Technometrics **35**, 403-410.

Handcock, M.S. and Wallis, J.R. (1994). *An Approach to Statistical Spatial–Temporal Modeling of Meteorological Fields* (with discussion) Journal of the American Statistical Association, 368-390.

Hoeksema, R.J. & Kitanidis, P.K. (1985). *Analysis of the Spatial Structure of Properties of Selected Aquifers.* Water Resources Research **21**, 563-572.

Intergovernmental Panel on Climate Change, (1990) *Scientific Assessment of Climate Change,* World Meteorological Organization, United Nations Environment Programme, Geneva, Switzerland.

Karl, T.R. & Williams, C.N., Jr., Quinlan, F.T. & Boden, T.A. (1990). *United States Historical Climatology Network (HCN) serial temperature and precipitation data.* NDP-019/R1, p. 83 plus appendices. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Kitanidis, P.K. (1983). *Statistical Estimation of Polynomial Generalized Covariance Functions and Hydrologic Applications.* Water Resources Research **19**, 909-921.

Kitanidis, P.K. & Lane, R.W. (1985). *Maximum Likelihood Parameter Estimation of Hydrologic Spatial Processes by the Gauss-Newton Method.* Journal of Hydrology **17**, 31-56.

Lettenmaier, D.P., Wood, E.F. & Wallis, J.R. (1992). *Hydrological Trends in the Continental United States.* To appear in Journal of Climate.

Mardia, K.V. & Marshall, R.J. (1984). *Maximum likelihood estimation of models for residual covariance in spatial regression.* Biometrika **71**, 135-146.

Matheron, G. (1965). *Les Variables Régionalisées et leur Estimation* Paris: Masson.

Matérn, B. (1986). *Spatial Variation* Second Ed, Lecture Notes in Statistics, **36.** Berlin: Springer-Verlag.

Mitchell, J. F.B. (1989). *The "Greenhouse Effect" and Climate Change.* Reviews of Geophysics **27**, 115-139.

Ripley, B.D. (1981). *Spatial Statistics.* New York: Wiley.

Wallis, J.R., Lettenmaier, D.P. & Wood, E.F. (1991). *A Daily Hydroclimatological Data Set for the Continental United States.* Water Resources Research **27**, 7, 1657-1663.

Whittle, P. (1954). *On Stationary Processes in the Plane.* Biometrika **41**, 434-449.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics.* Reprinted 1987, New York: Wiley.