Bayesian Multivariate Estimate of Global Temperature Trends

Dongchu Sun* Shawn Ni† Paul L. Speckman ‡

Abstract
In this paper, we consider a multivariate smoothing problem with correlated errors and correlated derivatives of the curves. Full Bayesian inference is introduced for the smoothing spline, the unknown covariance matrix of the errors and a symmetric smoothing parameter matrix. A prior is proposed on the symmetric smoothing parameter matrix, and Markov Chain Monte Carlo (MCMC) algorithms are developed for Bayesian computation. The proposed method is then applied to estimate the trends of abnormal surface temperature in ten geographical zones.

Key Words: Bayesian, Vector, Spline Smoothing

1. Introduction

Climate change is one of the most important subjects of scientific research of our time. Unbiased research on the causes and policy responses to global warming depend on accurate estimates of the trend of earth surface temperature. Estimates of the trend are based on temperature data measured in meteorological stations. The data contain measurement errors. Measurement errors stem from many sources such as inference of surface temperature change by human activities near meteorological stations and limitations in the measurement near polar regions. The data at different locations are also correlated (see Hansen and Lebedeff (1987)), which suggests correlated errors and trends. The trends are likely to be complex. Because of fluctuations in regional trends, annual surface temperatures generally do not move in lock step globally. For example, from 1940 to 1970 the temperature in the Northern Hemisphere decreased by about 0.5C (see Hansen et al., 1981). The complex trends of temperature in different locations may be modeled by a vector of smooth and correlated curves.

These features of the data suggest the need to develop a method of estimating multivariate trends that are correlated in the presence of potentially correlated errors. We develop a Bayesian approach to multivariate smoothing by smoothing splines. Suppose multivariate observations $\mathbf{y}_i = (y_{i1}, \ldots, y_{ip})$ are taken at points $t_1 < \cdots < t_n$, where $-\infty < a \le t_1$ and $t_n \le b < \infty$. Without loss of generality, we can assume $a = 0$ and $b = 1$. In the corresponding spline smoothing problem, a vector-valued unknown function $\mathbf{g}(s) = (g_1(s), \ldots, g_p(s))$ is chosen to minimize the loss function with a penalty on roughness,

$$\sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{g}(t_i))\mathbf{\Sigma}_0^{-1}(\mathbf{y}_i - \mathbf{g}(t_i))' + \int_0^1 \mathbf{g}^{(k)}(s)\mathbf{\Sigma}_1^{-1}(\mathbf{g}^{(k)}(s))'ds, \tag{1}$$

where $\mathbf{g}^{(k)}(s) = (g^{(1)}(s), \ldots, g^{(p)}(s))$. We can show that the minimizer of (1) is a vector-valued natural spline of degree $2k-1$.

*Department of Statistics, University of Missouri, Columbia, MO 65211, USA
†Department of Economics, University of Missouri, Columbia, MO 65211, USA
‡Department of Statistics, University of Missouri, Columbia, MO 65211, USA

In (1), $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$ are $p \times p$ positive definite penalty matrices on the approximation error and the "roughness" of $\mathbf{g}(t)$. Throughout the paper we also refer to them as covariance matrices. Using the notation "tr" for trace, the loss function (1) can be rewritten as

$$tr \left\{ \mathbf{\Sigma}_0^{-1} \left[\sum_{i=1}^{n} (\mathbf{y}_i - \mathbf{g}(t_i))'(\mathbf{y}_i - \mathbf{g}(t_i)) + \mathbf{\Sigma}_0 \mathbf{\Sigma}_1^{-1} \int_0^1 (\mathbf{g}^{(k)}(s))' \mathbf{g}^{(k)}(s) ds \right] \right\}. \tag{2}$$

When $p = 1$, the multivariate spline becomes a univariate smoothing spline, where the smooth component can be modeled as the univariate spline $g(t)$ that minimizes the loss function

$$\frac{1}{\sigma_0^2} \left\{ \sum_{i=1}^{n} (y_i - g(s_i))^2 + \frac{\sigma_0^2}{\sigma_1^2} \int_0^1 [g^{(k)}(s)]^2 ds \right\}. \tag{3}$$

In the smoothing spline literature, the noise-to-signal ratio $\eta = \sigma_0^2/\sigma_1^2$ is called the *smoothing parameter*. The smoothing parameter controls the balance between fidelity to the data and smoothness of the fitted function.

The problem of spline smoothing has been thoroughly studied for univariate models. See, for example, Wahba (1990) or Eubank (1999). Wahba (1985) and Wecker and Ansley (1983) showed that a univariate smoothing spline corresponds to a Bayesian linear mixed model and a state space model, respectively. Wang (1998) allowed for serial correlation in the errors in a univariate smoothing spline problem. A number of authors have considered restricted versions of multivariate smoothing splines with multivariate dependent variables, e.g, Yee and Wild (1996), Fessller (1991), and Wang et al. (2000). In these frequentist treatments of multivariate splines, the error covariance matrix $\mathbf{\Sigma}_0$ is unrestricted while $\mathbf{\Sigma}_1$ is restricted to be diagonal. These authors allow the penalty matrix $\mathbf{\Sigma}_0$ to be treated as either known (including the case where $\mathbf{\Sigma}_0$ is a function of i) or estimated as the covariance of residuals of univariate splines iteratively. In applications of multivariate models, the components of the multivariate function $\mathbf{g}(t)$ may be influenced by common factors. For instance, temperature of different zones have correlated trends. Restricting $\mathbf{\Sigma}_1$ to be diagonal is inadvisable under such a scenario. In this study we treat both $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$ as unknown and perform joint Bayesian inference on the covariance matrices and $\mathbf{g}(t)$.

To our knowledge, the multivariate smoothing spline has not been treated in this generality. Our contributions are as follows. First, we derive the minimizer of (1). Second, we note that the multivariate analog of the univariate smoothing parameter η is $\mathbf{\Sigma}_0\mathbf{\Sigma}_1^{-1}$, which is not symmetric and is overparameterized. We define a suitable multivariate smoothing parameter. Third, it is well known that the univariate smoothing spline is equivalent to the Bayes estimate with a generalized Gaussian prior (Kimeldorf and Wahba (1970).) There is also a Bayesian interpretation to the multivariate case, which we exploit with a fully Bayesian analysis. Informative priors are justifiable in some cases, but they may not be appropriate in all applications. When the researcher has limited knowledge on the smoothness of the spline, objective (non-informative) priors become useful. However, objective priors may render the posterior improper in univariate linear mixed models (see e.g., Hill, 1965, and Hobert and Casella, 1996). For the univariate case, Sun et al. (1999) and Speckman and Sun (2003) derived conditions on the priors of $(\sigma_0^2, \sigma_0^2/\sigma_1^2)$ that give rise to a proper posterior for smoothing splines. For the general case with p greater than one, analysis of objective priors is scant. The present study seeks to fill this void.

2. Multivariate Spline Smoothing

Problem (2) can be shown to be equivalent to finding a matrix $\mathbf{Z}$ that solves

$$\min_{\mathbf{Z}} \mathrm{tr}\left\{\mathbf{\Sigma}_0^{-1}(\mathbf{Y}-\mathbf{Z})'(\mathbf{Y}-\mathbf{Z}) + \mathbf{\Sigma}_1^{-1}\mathbf{Z}'\mathbf{Q}\mathbf{Z}\right\}. \tag{4}$$

Here $\mathbf{Q}$ is the positive definite matrix of rank $n-k$ from the univariate problem.

Now let $\mathbf{y} = vec(\mathbf{Y})$ and $\mathbf{z} = vec(\mathbf{Z})$. Using the fact that $tr(\mathbf{ABCD}) = vec'(\mathbf{D})(\mathbf{A}\otimes\mathbf{C}')vec(\mathbf{B}')$ for any conforming matrices $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$, (4) is equivalent to

$$\min_{\mathbf{z}} \Big\{(\mathbf{y}-\mathbf{z})'(\mathbf{\Sigma}_0^{-1}\otimes\mathbf{I}_n)(\mathbf{y}-\mathbf{z}) + \mathbf{z}'(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{Q})\mathbf{z}\Big\}. \tag{5}$$

The solution to (5) is

$$\hat{\mathbf{z}} = (\mathbf{I}_{np} + \mathbf{\Sigma}_0\mathbf{\Sigma}_1^{-1}\otimes\mathbf{Q})^{-1}\mathbf{y}. \tag{6}$$

2.1 The multivariate smoothing parameter

One central issue in defining the multivariate smoothing spline is to generalize the smoothing parameter η when $p=1$ in (3) to the general case, where the analog is the matrix $\mathbf{\Sigma}_0\mathbf{\Sigma}_1^{-1}$ in (2). However, $\mathbf{\Sigma}_0\mathbf{\Sigma}_1^{-1}$ is not an ideal smoothing parameter matrix because it is not symmetric and it is overparameterized with p^2 parameters. A matrix version of the smoothing parameter should be symmetric with $p(p+1)/2$ free parameters. We reparameterize $(\mathbf{\Sigma}_0, \mathbf{\Sigma}_1)$ as

$$\mathbf{\Sigma}_0^{-1} = \mathbf{\Psi}'\mathbf{\Psi}, \tag{7}$$

$$\mathbf{\Sigma}_1^{-1} = \mathbf{\Psi}'\Xi\mathbf{\Psi}, \tag{8}$$

where $\mathbf{\Psi}$ is a $p\times p$ invertible matrix with $p(p+1)/2$ free parameters and Ξ is symmetric. The $p\times p$ positive definite matrix Ξ is a matrix version of the noise-to-signal ratio or smoothing parameter. When $p=1$, Ξ is exactly the smoothing parameter σ_0^2/σ_1^2. For $p>1$, decompositions (7) and (8) imply $\Xi = \mathbf{\Psi}^{-T}\mathbf{\Sigma}_1^{-1}\mathbf{\Psi}^{-1}$ and $\mathbf{\Sigma}_0\mathbf{\Sigma}_1^{-1} = \mathbf{\Psi}^{-1}\Xi\mathbf{\Psi}$. With this definition, solution (6) becomes

$$\begin{aligned}\hat{\mathbf{z}} &= (\mathbf{I}_{np} + \mathbf{\Psi}^{-1}\Xi\mathbf{\Psi}\otimes\mathbf{Q})^{-1}\mathbf{y}\\ &= (\mathbf{\Psi}^{-1}\otimes\mathbf{I}_n)(\mathbf{I}_{np}+\Xi\otimes\mathbf{Q})^{-1}(\mathbf{\Psi}\otimes\mathbf{I}_n)\mathbf{y}.\end{aligned} \tag{9}$$

2.2 Bayes Estimates of $\mathbf{z}$ for Fixed $(\mathbf{\Sigma}_0, \mathbf{\Sigma}_1)$

Consider the model

$$\mathbf{y}_i = \mathbf{z}_i + \boldsymbol{\epsilon}_i, \tag{10}$$

where $\boldsymbol{\epsilon}_i' \sim N(0, \mathbf{\Sigma}_0)$. The density (likelihood) of $\mathbf{y}$ given $\mathbf{z}$ and $\mathbf{\Sigma}_0$ based on model (10) is

$$f(\mathbf{y}\mid\mathbf{z},\mathbf{\Sigma}_0) = \frac{1}{(2\pi)^{\frac{np}{2}}|\mathbf{\Sigma}_0|^{\frac{n}{2}}}\exp\Big\{-\frac{1}{2}(\mathbf{y}-\mathbf{z})'(\mathbf{\Sigma}_0^{-1}\otimes\mathbf{I}_n)(\mathbf{y}-\mathbf{z})\Big\}. \tag{11}$$

Motivated by (5), we assume the partially informative normal distribution (Sun et al., 1999) for $\mathbf{z}$ since $\mathbf{Q}$ typically does not have full rank,

$$f(\mathbf{z}\mid\mathbf{\Sigma}_1) \propto \Big|\frac{1}{2\pi}\Big(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{Q}\Big)\Big|_+^{1/2}\exp\Big\{-\frac{1}{2}\mathbf{z}'(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{Q})\mathbf{z}\Big\}, \tag{12}$$

where $|\mathbf{A}|_+$ is the product of positive eigenvalues of a nonnegative definite matrix $\mathbf{A}$. The following fact is immediate and its proof is omitted.

Fact 1 *For fixed* $(\mathbf{\Sigma}_0, \mathbf{\Sigma}_1)$*, the conditional posterior of* $\boldsymbol{z}$ *given* $\boldsymbol{y}$ *is*

$$(\boldsymbol{z} \mid \boldsymbol{y}, \mathbf{\Sigma}_0, \mathbf{\Sigma}_1) \quad \sim \quad N_{pn}(\hat{\boldsymbol{z}}, \boldsymbol{M}), \tag{13}$$

where $\hat{\boldsymbol{z}}$ *is given by (6) and*

$$\boldsymbol{M} = (\mathbf{\Sigma}_0^{-1} \otimes \boldsymbol{I}_n + \mathbf{\Sigma}_1^{-1} \otimes \boldsymbol{Q})^{-1}. \tag{14}$$

Remark 1 For fixed $(\mathbf{\Sigma}_0, \mathbf{\Sigma}_1)$, the solution of smoothing spline (6) coincides with the posterior mean of $\mathbf{z}$ under the prior (12).

3. Priors for $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$

A key aspect of Bayesian analysis is the choice of prior. For a full Bayesian analysis, researchers must choose priors for the variances. Using the special case with $p = 1$ that independent priors on σ_0^2 and σ_1^2 are undesirable in the context of smoothing splines.

As an alternative to independent priors on the variances, one can specify a prior on $\eta = \sigma_0^2/\sigma_1^2$. For example, if a researcher is ignorant on the relative size of the variances, then it is reasonable to assume $\xi = \frac{\eta}{1+\eta}$ is uniform on (0,1). One may assign a corresponding prior on η and an independent prior on σ_0^2. This will avoid choosing priors that have unintendedly strong influence on the variance ratio.

When $p = 1$, Koop and Van Dijk (2000) suggest the prior

$$\pi(\eta) = \frac{1}{(\eta+1)^2}, \quad \eta > 0. \tag{15}$$

This is equivalent to assuming a Uniform $(0, 1)$ prior for $\sigma_1^2/(\sigma_0^2 + \sigma_1^2)$. It is easy to see that the noise-to-signal ratio $\eta = \sigma_0^2/\sigma_1^2$, and assume $\pi(\sigma_1^2/(\sigma_0^2 + \sigma_1^2)) = \text{Uniform}(0, 1)$. Then the prior for $\sigma_1^2/(\sigma_0^2 + \sigma_1^2)$ and $\sigma_0^2/(\sigma_0^2 + \sigma_1^2)$ are then $\pi(\frac{1}{1+\eta})$ and $\pi(\frac{1}{1+\frac{1}{\eta}})$ are Uniform$(0, 1)$. Then the prior distribution of η is $\pi(\eta) \propto \frac{1}{(\eta+1)^2}$. Moreover, the prior distribution of $\frac{1}{\eta}$ is $\pi(\frac{1}{\eta}) \propto \frac{1}{(1+\frac{1}{\eta})^2}$, which takes the same form as that of η.

A common prior for a $p \times p$ covariance matrix $\mathbf{\Sigma}$ is an Inverse Wishart type prior, $\text{IW}_p(m, \mathbf{A})$, with density

$$[\text{IW}(\mathbf{\Sigma} \mid m, \mathbf{A})] \propto |\mathbf{\Sigma}|^{-\frac{m+p+1}{2}} etr(-\frac{1}{2}\mathbf{\Sigma}^{-1}\mathbf{A}), \tag{16}$$

where $etr(\cdot)$ stands for $\exp[\text{tr}(\cdot)]$, m is often interpreted as degrees of freedom and $\mathbf{A}$ is a known nonnegative definite matrix. The discussion based on the univariate model suggests that assigning independent priors on $\mathbf{\Sigma}_0$ and $\mathbf{\Sigma}_1$ of the type (16) renders unintended consequences on the matrix version of the noise-signal ratio.

Instead, we propose the following prior density for the multivariate noise-to-signal ratio Ξ,

$$\pi(\Xi \mid b) = \frac{b^{\frac{(p+1)p}{2}} \Gamma_p(\frac{2(p+1)}{2})}{\left(\Gamma_p(\frac{p+1}{2})\right)^2} |\Xi + b\mathbf{I}_p|^{-(p+1)}, \tag{17}$$

where $b > 0$ is a scale parameter and $\Gamma_p(\frac{n}{2}) = \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \Gamma(\frac{n}{2} - \frac{j-1}{2})$ for any $n > p$. The prior (17) has two interesting features. First, the prior can be simulated through a hierarchical structure. Throughout the paper, we use "( )" for distribution and "[ ]" for density.

When $b = 1$ in prior (17), despite the implied symmetry in the form of priors of 'noise-signal ratio' and 'signal-noise ratio', the prior is actually quite informative. The penalty in (1) depends crucially on the scaling of t, which we assumed to be in $[0, 1]$. It follows that the prior on η in the univariate case or Ξ in the multivariate case must also have a scale factor. Unfortunately, it seems difficult to elicit a prior on the variance of $g^{(k)}$, even in the univariate case.

The solution adopted by White (2006) is to elicit a prior in terms of the effective degrees of freedom of the smoother. From (9), the smoother matrix for the multivariate problem is

$$S(\Xi) = (\mathbf{\Psi}^{-1} \otimes \mathbf{I}_n)(\mathbf{I}_{np} + \Xi \otimes \mathbf{Q})^{-1}(\mathbf{\Psi} \otimes \mathbf{I}_n).$$

We define the effective degrees of freedom for a nonparametric linear smoother of the form $\mathbf{Sy}$ as

$$tr(S(\Xi)) = tr((\mathbf{I}_{np} + \Xi \otimes \mathbf{Q})^{-1}).$$

Following White (2006), we choose b so that the median of the distribution of $tr(S(\Xi))$ under prior (17) is consistent with prior belief on the complexity of the curves to be fitted. Complexity can be envisioned as the number of parametric terms needed to fit the curve in a regression model. Of course, the complexity depends on the amount of noise in the data as well as the number of observations. Less noise or more observations will admit a more complex fit. In practice, we estimate the median of the prior effective degrees of freedom by Monte Carlo simulation using a hierarchical scheme.

When $p = 1$ the univariate version of prior (17) becomes

$$\pi(\eta) = \frac{b}{(b + \eta)^2}. \tag{18}$$

The hyperparameter b can be selected in the univariate case based on the same principle as in the multivariate case.

Our numerical simulations show that if b is the hyperparameter for univariate spline, a reasonable hyperparameter for the corresponding p-dimensional multivariate spline is pb.

Table 1: Median of Trace of the Smoother tr($S(\Xi)$) with Hyper-Parameter pb in Prior (17)

b	1	100	1000	2000
$p = 1$	53.7	17.8	10.4	8.9
$p = 2$	96.9	32.2	18.8	16.3
$p = 3$	136.3	45.5	26.7	23.1
$p = 4$	173.7	57.8	34.5	29.5

n=150. Ξ is generated from distribution (17).

For univariate cubic smoothing splines with $t_i = i$, Shiller(1984) showed that the matrix $\mathbf{Q}$ in (4) is $\mathbf{Q} = \mathbf{F}_0'\mathbf{F}_1^{-1}\mathbf{F}_0$, where

$$\mathbf{F}_0 = \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{pmatrix}_{(T-2)\times T} \quad \text{and } \mathbf{F}_1 = \frac{1}{6}\begin{pmatrix} 4 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 4 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & 1 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{pmatrix}_{(T-2)\times(T-2)}.$$

The proposed prior for $\mathbf{\Sigma}_0$ is type (16)

$$\pi(\mathbf{\Sigma}_0) \quad \sim \quad \text{IW}_p(m_0, \mathbf{Q}_0) \propto |\mathbf{\Sigma}_0|^{-\frac{m_0+p+1}{2}} etr(-\frac{1}{2}\mathbf{\Sigma}_0^{-1}\mathbf{Q}_0).$$

If $m_0 = p + 1$ and $\mathbf{Q}_0^{-1} \to \mathbf{0}$, the prior for $\mathbf{\Psi}$ approaches the right Haar (RH) prior

$$\pi_{RH}(\mathbf{\Psi}) \quad \propto \quad \prod_{i=1}^{p} \frac{1}{\psi_{ii}^{i}}, \tag{19}$$

where ψ_{ii} is the ith diagonal element of $\mathbf{\Psi}$. For an iid normal $(\boldsymbol{\mu}, \mathbf{\Sigma}_0)$ population, Berger and Sun (2008) showed that this right Haar prior is a matching prior (meaning the posterior credible intervals is the same as frequentist confidence interval of the same level). We propose an independent RH prior (19) for $\mathbf{\Sigma}_0$ and prior (17) for Ξ.

In the case of the univariate model $p = 1$, (19) is equivalent to

$$\pi(\sigma_0^2) \propto \frac{1}{\sigma_0^2}, \tag{20}$$

which is also the Jeffreys type prior for the univariate case.

4. Application: Estimate the Trends in Earth Surface Temperature Using Multivariate Splines

To remove seasonal variations within a year, annual averages of abnormal temperatures from 1880 to 2012 are used for estimation. The data are 1880-2012 annual mean temperature anomalies in degrees Celsius in 10 zones: Zone 1: 24N-90N, Zone 2: 24S-24N, Zone 3: 90S-24S, Zone 4: 64N-90N, Zone 5: 44N-64N, Zone 6: 24N-44N, Zone 7: Equator-24N, Zone 8: 24S-Equator, Zone 9: 44S-24S, Zone 10: 64S-44S. The data are obtained from the website of NASA's Goddard Institute for Space Studies ($http : //data.giss.nasa.gov/gistemp/tabledata_v3/ZonAnn.Ts+dSST.txt$).

Figure 1 plots the data of the ten time series and the smoothed trends. For each zone, the same color is used for the upper and lower panel in the figure.

The estimated trends of temperature anomalies in different zones are quite smooth and show strong co-movements. There is a considerable difference in the magnitude of the trends. The estimated trend in Zone 4 (blue dot-dashed line) increases sharply from 1880 to 1940s, then starts to decline from 1940s to 1970s, and rises again from 1970s to 2012. The last stage of the increasing trend is approximately convex in time, which indicates positive acceleration. To a lesser extent, the estimated trends of Zone 5 (light blue dashed line) and Zone 1 (black solid line) also exhibit the three-stage "increase-decrease-increase" pattern, and the timing of the turning points coincide with those of Zone 4 trends. The estimated trends of the remaining seven zones are very similar. All are flat from 1880 to 1970s and then start a moderate rise.

We now turn to the standard deviation-correlation matrix (i. e., a matrix with standard deviations on the diagonal and correlations on the off-diagonals) of the residuals of the ten zones.

$$\begin{pmatrix} .170 & .351 & .242 & .641 & .911 & .648 & .336 & .308 & .347 & .057 \\ .351 & .148 & .231 & .118 & .205 & .495 & .946 & .923 & .511 & -.193 \\ .242 & .231 & .117 & .222 & .158 & .193 & .186 & .259 & .598 & .599 \\ .641 & .118 & .222 & .374 & .484 & .004 & .053 & .169 & .166 & .147 \\ .911 & .205 & .158 & .484 & .251 & .453 & .178 & .190 & .261 & .004 \\ .648 & .495 & .193 & .004 & .453 & .150 & .549 & .372 & .367 & .013 \\ .336 & .946 & .186 & .053 & .178 & .549 & .145 & .758 & .418 & -.149 \\ .308 & .923 & .259 & .169 & .190 & .372 & .758 & .183 & .550 & -.204 \\ .347 & .511 & .598 & .166 & .261 & .367 & .418 & .550 & .123 & -.056 \\ .057 & -.193 & .599 & .147 & .004 & .013 & -.149 & -.204 & -.056 & .218 \end{pmatrix}.$$

The residuals of the fitted curves are correlated. The correlations are mostly positive but vary in magnitude. The strongest correlations are between Zone 1 and Zone 4, Zone 1 and Zone 5, and Zone 1 and Zone 6. Some correlations are negative. The largest negative correlation is between Zone 8 and Zone 10.

The following is the posterior mean of $\mathbf{\Sigma}_1$ multiplied by 30000.

$$\begin{pmatrix} 1.02 & 0.27 & -0.12 & 1.38 & 1.08 & 0.84 & 0.51 & -0.06 & 0.00 & -0.15 \\ 0.27 & 0.27 & 0.06 & 0.45 & 0.24 & 0.24 & 0.24 & 0.33 & 0.15 & -0.03 \\ -0.12 & 0.06 & 0.36 & 0.09 & -0.12 & -0.18 & -0.18 & 0.42 & 0.30 & 0.42 \\ 1.38 & 0.45 & 0.09 & 3.36 & 1.26 & 0.75 & 0.36 & 0.57 & 0.27 & -0.12 \\ 1.08 & 0.24 & -0.12 & 1.26 & 1.29 & 0.84 & 0.48 & -0.12 & -0.03 & -0.12 \\ 0.84 & 0.24 & -0.18 & 0.75 & 0.84 & 0.87 & 0.57 & -0.24 & -0.09 & -0.21 \\ 0.51 & 0.24 & -0.18 & 0.36 & 0.48 & 0.57 & 0.54 & -0.15 & -0.09 & -0.24 \\ -0.06 & 0.33 & 0.42 & 0.57 & -0.12 & -0.24 & -0.15 & 1.02 & 0.48 & 0.27 \\ 0.00 & 0.15 & 0.30 & 0.27 & -0.03 & -0.09 & -0.09 & 0.48 & 0.36 & 0.24 \\ -0.15 & -0.03 & 0.42 & -0.12 & -0.12 & -0.21 & -0.24 & 0.27 & 0.24 & 0.72 \end{pmatrix}.$$

The second order derivatives of the curves are correlated. The strongest positive correlations are between Zone 1 and Zone 4, Zone 1 and Zone 5, and Zone 1 and Zone 6. The same zone pairs that also have strong correlations of the residuals.

The correlation in the residuals and the co-movement in the trends suggest the multivariate approach is appropriate for estimating the trends in surface temperature.

REFERENCES

Berger, J. O. and Sun, D. (2008), 'Objective priors for the bivariate normal model', *Annals of Statistics* **36**, 963–982.

Eubank, R. L. (1999), *Nonparametric Regression and Spline Smoothing*, Marcel Dekker Inc.

Fessler, J. A. (1991), 'Nonparametric fixed-interval smoothing with vector splines', *IEEE Transactions on Acoustics, Speech, and Signal Processing* **39**, 852–859.

Hansen, J., D. Johnson, A. Lacis, S. Lebedeff, P. Lee, D. Rind, and G. Russell, 1981: Climate impact of increasing atmospheric carbon dioxide. *Science*, **213**, 957-966.

Hansen, J.E., and S. Lebedeff, (1987), 'Global trends of measured surface air temperature', *J. of Geophys. Res.*, **92**, 13345-13372.

Hill, B. M. (1965), 'Inference about variance components in the one-way model', *Journal of the American Statistical Association* **60**, 806–825.

Hobert, J. and Casella, G. (1996), 'The effect of improper priors on gibbs sampling in hierarchicallinear mixed models', *Journal of the American Statistical Association* **91**, 1461–1473.

Kimeldorf, George S. and Wahba, Grace (1970), "A Correspondence Between Bayesian Estimation on Stochastic Processes and Smoothing By Splines," *The Annals of Mathematical Statistics*, 41, 495–502.

Koop, Gary and van Dijk, Herman K. (2000), "Testing for integration using evolving trend and seasonals models: A Bayesian approach," *Journal of Econometrics*, 97, 261–291.

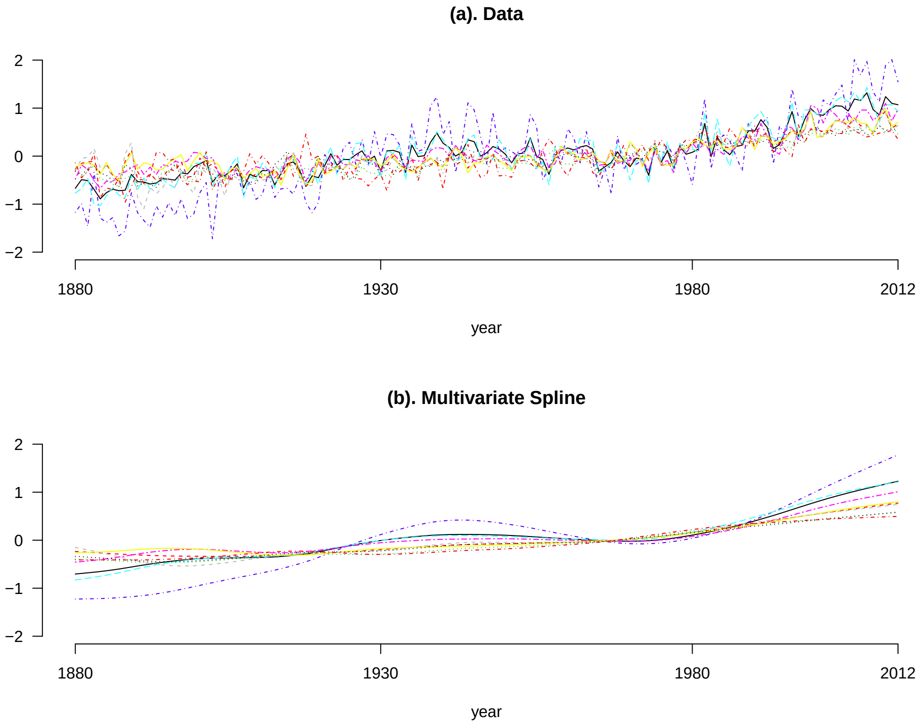

Figure 1: Data and Estimated Trends of Temperature Anomalies in Degrees Celsius in 10 Zones

Shiller, R. J. (1984), 'Smoothness priors and nonlinear regression', *Journal of the American Statistical Association* **79**, 609–615.

Speckman, P. L. and Sun, D. (2003), 'Full (b)ayesian spline smoothing and intrinsic autoregressive priors', *Biometrika* **90**, 289–302.

Sun, D., Tsutakawa, R. K. and Speckman, P. L. (1999), 'Posterior distribution of hierarchical models using CAR(1) distributions', *Biometrika* **86**, 341–350.

Wahba, G. (1985), 'A comparison of Gcv and Gml for choosing the smoothing parameter in the generalized spline smoothing problem', *The Annals of Statistics* **13**, 1378–1402.

Wahba, G. (1990), *Spline Models for Observational Data*, Philadelphia: Society for Industrial and Applied Mathematics.

Wang, Y. (1998), 'Smoothing spline models with correlated random errors', *Journal of the American Statistical Association* **93**, 341–348.

Wang, Y., Guo, W. and Brown, M. B. (2000), 'Spline smoothing for bivariate data with applications to association between hormones', *Statistica Sinica* **10**(2), 377–397.

Wecker, W. E. and Ansley, C. F. (1983), 'The signal extraction approach to nonlinear regression and spline smoothing', *Journal of the American Statistical Association* **78**, 81–89.

Yee, T. W. and Wild, C. J. (1996), 'Vector generalized additive models', *Journal of the Royal Statistical Society, Series B: Methodological* **58**, 481–493.