Spatio-temporal modeling for real-time ozone forecasting

Lucia Paci[a,1,*], **Alan E. Gelfand**[a], and **David M. Holland**[b]

[a]Department of Statistical Science at Duke University, Box 90251, Durham NC 27708-0251, USA

[b]U.S. Environmental Protection Agency, National Exposure Research Laboratory, Research Triangle Park, NC 27711, USA

Abstract

The accurate assessment of exposure to ambient ozone concentrations is important for informing the public and pollution monitoring agencies about ozone levels that may lead to adverse health effects. High-resolution air quality information can offer significant health benefits by leading to improved environmental decisions. A practical challenge facing the U.S. Environmental Protection Agency (USEPA) is to provide real-time forecasting of current 8-hour average ozone exposure over the entire conterminous United States. Such *real-time* forecasting is now provided as spatial forecast maps of *current* 8-hour average ozone defined as the average of the previous four hours, current hour, and *predictions* for the next three hours. Current 8-hour average patterns are updated hourly throughout the day on the EPA-AIRNow web site.

The contribution here is to show how we can substantially improve upon current real-time forecasting systems. To enable such forecasting, we introduce a downscaler fusion model based on first differences of real-time monitoring data and numerical model output. The model has a flexible coefficient structure and uses an efficient computational strategy to fit model parameters. Our hybrid computational strategy blends continuous background updated model fitting with real-time predictions. Model validation analyses show that we are achieving very accurate and precise ozone forecasts.

Keywords

data fusion; hierarchical model; kriging; Markov chain Monte Carlo; space-time covariance; time differencing

1. Introduction

The evaluation and control of air pollution levels are fundamental environmental issues for environmental decision-makers. The United States Environmental Protection Agency (USEPA) developed the AIRNow web site (http://www.airnow.gov) to provide the public, air regulatory agencies and health scientists with easy access to *real-time* national air pollution information. Current and next day forecasts of ozone and fine particulate matter are produced at over 300 cities across the United States (U.S.) on a daily basis. For ozone, forecasts at these monitoring sites are then interpolated across the continent, at a chosen spatial scale, to provide forecast maps for current 8-hour average ozone levels and next day patterns of 8-hour maximum ozone concentration. We focus here on *current* 8-hour average

---
*Corresponding author. Tel.: +39 051 2098201; Fax: +39 051 232153. lucia.paci2@unibo.it (Lucia Paci), alan@stat.duke.edu (Alan E. Gelfand), holland.david@epa.gov (David M. Holland).
[1]Permanent address: Department of Statistical Science at the University of Bologna, Via delle Belle Arti, 41 - 40126 Bologna, Italy.

patterns which are updated hourly throughout the day on the AIRNow web site in the form of point estimates with no associated uncertainties.

Measurements at monitoring stations present the most direct and accurate way to obtain air quality information. However, monitoring sites are often sparsely and irregularly spaced over large areas and affected by missingness. These data are the sole data source used to develop the AIRNow forecasts. However, a second source of real-time spatial information is available that could be used to improved forecasting. A numerical atmospheric model known as the Eta-Community Multi-Scale Air Quality (CMAQ) model (http://www.epa.gov/asmdnerl/CMAQ) provides predictions of average pollution concentrations at the 12 km grid cell resolution for successive time periods including 48 hours into the future. At this resolution, we have hourly numerical model information for approximately 54, 000 grid cells spanning the conterminous U.S. However, these predictions are expected to be biased with unknown calibration.

The contribution of this work is to develop a space-time data assimilation strategy to enable use of both data sources to provide the forecasts of current 8-hour average ozone level in *real-time*. Here, current 8-hour ozone is defined as the average of the previous four hours, current hour, and *predictions* for the next three hours. We combine data from the real-time ozone monitoring network with the output from the Eta-CMAQ computer model, using first differences along with a linear regression model having spatio-temporally varying coefficients. We propose a combination of offline fitting and online prediction to enable feasible real-time forecasting. In particular, we obtain a fully model-based fusion using a Bayesian hierarchical model. However, we implement spatial kriging and temporal forecasting to 54,000 locations using predictive output under this model with an available kriging package. Our hybrid strategy enables uncertainty assessment associated with these predictions. To evaluate the accuracy of the forecasts, we use historical data to show that our overall approach validates well and provides significant improvement in the accuracy of forecasting relative to that of AIRNow. Again, we do not attempt to identify and implement a "best" model. Rather, we provide a model-based strategy that can be run in real time while improving on the existing forecasting as well as adding uncertainty.

Sahu et al. [1] proposed an alternative approach for forecasting based on a Bayesian spatio-temporal model applied to hourly ozone concentrations. They used data over a running window of seven days to predict 8-hour average ozone level for the current hour. To allow real-time hourly forecasting, they developed a spatial regression model that avoids iterative algorithms such as Markov Chain Monte Carlo (MCMC) methods. In Sahu et al. [2], a dynamic model is developed for forecasting next day 8-hour maximum ozone patterns. However, the dynamic model is computationally intensive and not feasible for use in real-time forecast applications. Kang et al. [3] consider Kalmanfilter approaches to improve next day forecasts of ozone concentrations at individual U.S. monitoring sites for the summer of 2005.

Recently, several papers have appeared in the literature considering data fusion methods for combining observed data and computer model output (see Gelfand and Sahu [4] and references therein). Fuentes and Raftery [5] present an application of the Bayesian melding approach [6] where observed data are combined with numerical model output by introducing a 'true' latent point-level process driving both sources of data. Model output is expressed as integrals over a grid cells of the latent point-level process while monitoring data are linked to the latent process via measurement error model. A spatio-temporal extention of Bayesian melding has been proposed by McMillan et al. [7] for modeling fine particulate matter.

Berrocal et al. [8] and Berrocal et al. [9] developed univariate and bivariate downscaler models that relate the monitoring data and the numerical model output using a regression model with spatially varying coefficients [10] modeled, in turn, as Gaussian processes. This approach avoids the change-of-support problem which arises in previous hierarchical models for data fusion settings to combine point- and grid-referenced data. The approach accommodates local calibration of the numerical model output, avoiding computational limitations of the Bayesian melding. Berrocal et al. [11] proposed two neighbor-based extensions of the downscaler model. The first extension introduces a Gaussian Markov random field to smooth the computer model output, while the second extension introduces spatially varying weights driven by a latent Gaussian process. Unfortunately, full implementation of these modeling approaches again, is infeasible for real-time forecasting.

The monitoring data we use is collected from 717 real-time monitoring stations operating around eastern U.S., for the two-week period August 1-14, 2011. The Eta-CMAQ model output includes 21,109 grid cells spanning our illustrative study region. In practice, real-time hourly output from the Eta-CMAQ model is available up to 48 hours in the future.

The paper is organized as follows. In Section 2 we describe the data used in this study. In Section 3, we review the univariate downscaler and we present our strategy to produce real-time 8-hour average ozone forecasts. Model fitting details are discussed in Section 4, with computational details deferred to an appendix. Our prediction method is developed in Section 5 while Section 6 provides the analyses and results, including comparison with the above work of Sahu et al. [1] and with AIRNow predictions. A brief Section 7 offers conclusions and future work.

2. Data description

The first source of data consists of current 8-hour average ozone concentrations in parts per billion units (ppb) collected at 717 real-time monitoring stations operating in the eastern U.S. during a two-week period over August 1-14, 2011; see Figure 1. The region used in our application covers roughly half the conterminous U.S. and the monitoring sites farthest apart are about 2860 km from each other. We set aside data from 70 monitoring sites for validation purposes; these sites were chosen at random (again, see Figure 1).

The second source of data is the numerical output of the Eta-CMAQ model. This model uses meteorological information, emission inventories, and land usage to estimate average pollution levels for gridded cells (at 12 km^2 resolution) over successive time periods without any missing values. There are 21,109 Eta-CMAQ grid cells spanning our study region. In practice, real-time hourly output from the Eta-CMAQ model is available up to 48 hours in the future.

3. Modeling

We briefly review the univariate spatio-temporal downscaler presented in Berrocal et al. [8] and we offer a proposed model for current 8-hour average ozone concentration. Then, we present our strategy to obtain real-time and accurate predictions within the real-time environment.

3.1. Downscaler for 8-hour average ozone level

Let $Y_t(\mathbf{s})$ denote the hourly ozone concentration at a generic location $\mathbf{s}$ and $W_t(B)$ be the Eta-CMAQ output over grid cell B. The downscaler model addresses the difference in spatial resolution between monitoring data and numerical model output, by associating to each site

s the grid cell B that contains $\mathbf{s}$. Then, the model links the observational data and the Eta-CMAQ output as follows:

$$Y_t(\mathbf{s}) = \tilde{\beta}_{0,t}(\mathbf{s}) + \tilde{\beta}_{1,t}(\mathbf{s}) W_t(B) + \epsilon_t(\mathbf{s}) \quad (1)$$

where

$$\begin{aligned} \tilde{\beta}_{0,t}(\mathbf{s}) &= \beta_0 + \beta_{0,t}(\mathbf{s}) \\ \tilde{\beta}_{1,t}(\mathbf{s}) &= \beta_1 + \beta_{1,t}(\mathbf{s}) \end{aligned} \quad (2)$$

and $\epsilon_t(\mathbf{s})$ is a white noise process with nugget variance τ^2.[2] $\tilde{\beta}_{0,t}(\mathbf{s})$ and $\tilde{\beta}_{1,t}(\mathbf{s})$ can be interpreted as a spatio-temporal intercept process and a spatio-temporal slope process, respectively. Equivalently, $\beta_{0,t}(\mathbf{s})$ and $\beta_{1,t}(\mathbf{s})$ in (2) can be viewed as local spatio-temporal adjustment to the overall intercept β_0 and global slope β_1.

Now, consider the current 8-hour average ozone level $Z_t(\mathbf{s})$ defined, from above, as the average of the previous four hours, the current hour and the next three hours in the future, that is

$$Z_t(\mathbf{s}) = \frac{1}{8} \sum_{k=-4}^{+3} Y_{t+k}(\mathbf{s}) \quad (3)$$

According to the definition in (3), with the model in (1)-(2), the downscaler model for $Z_t(\mathbf{s})$ is given by:

$$Z_t(\mathbf{s}) = \frac{1}{8} \sum_{k=-4}^{+3} \tilde{\beta}_{0,t+k}(\mathbf{s}) + \frac{1}{8} \sum_{k=-4}^{+3} \tilde{\beta}_{1,t+k}(\mathbf{s}) W_{t+k}(B) + \frac{1}{8} \sum_{k=-4}^{+3} \epsilon_{t+k}(\mathbf{s}) \quad (4)$$

Modeling the 8-hour averages, $Z_t(\mathbf{s})$ in (4), will be not feasible within a real-time environment. The induced dependence structure in the $Z_t(\mathbf{s})$ process will become very messy and intractable for fast model fitting; consider, for example, the induced association between $Z_t(\mathbf{s})$ and $Z_{t-1}(\mathbf{s}')$. However, if we work with differences we can simplify the specifications and can still capture the ozone diurnal variation, the influence of the Eta-CMAQ output, and the space-time random variation. Moreover, less uncertainty is associated to the predictions when modeling monitoring data differences compared with modeling the hourly ozone concentrations and converting to the $Z_t(\mathbf{s})$'s, as we will see below.

In fact, with the objective of expediting computation, we will also simplify the models for $\beta_{0,t}(\mathbf{s})$ and $\beta_{1,t}(\mathbf{s})$ to multiplicative forms in space and time. With say M locations and T time points, we reduce from $2MT$ to $2(M+T)$ latent variables. As we clarify below, this will not imply space-time separability of dependence structure. So, let

$$\begin{aligned} \beta_{0,t}(\mathbf{s}) &= \beta_0(\mathbf{s}) \beta_{0,t} \\ \beta_{1,t}(\mathbf{s}) &= \beta_1(\mathbf{s}) \beta_{1,t} \end{aligned} \quad (5)$$

where $\beta_0(\mathbf{s})$ and $\beta_1(\mathbf{s})$ represent the pure spatial effects while $\beta_{0,t}$ and $\beta_{1,t}$ denote the pure temporal components. We note that the components of these products are only identifiable up to scale; in modeling we will only introduce a variance component for the spatial process. We can also model $\beta_{0,t}(\mathbf{s})$ and $\beta_{1,t}(\mathbf{s})$, respectively, in an additive fashion, as a sum of

---

[2]In principle, other explanatory variables, such as real-time temperature or elevation, could be added to the downscaler model. Moreover, these variables can be at areal or point scale. However, here we confine ourselves to the computer model output.

temporal and spatial components. However, the multiplicative form in (5) leads to more flexible local adjustment, as we clarify below.

3.2. Downscaler for monitoring data differences

Denote the monitoring data differences $\Delta_t^Z(\mathbf{s})$ by

$$\Delta_t^Z(\mathbf{s}) = 8\left(Z_t(\mathbf{s}) - Z_{t-1}(\mathbf{s})\right). \quad (6)$$

First differences are a commonly-used tool in time series analysis settings and motivate the introduction of $\Delta_t^Z(\mathbf{s})$. The spatial time series of first differences in (6) is more stable than the original series and enable us to highlight the short-term pattern which strongly characterizes the ozone levels. Moreover, we can reduce our attention from eight to only two elements when we compute monitoring data differences. That is, we have

$$\Delta_t^Z(\mathbf{s}) = Y_{t+3}(\mathbf{s}) - Y_{t-5}(\mathbf{s}). \quad (7)$$

Suppose we insert (1) into (7). For the resulting $\Delta_t^Z(\mathbf{s})$, the overall intercept β_0 will disappear and we obtain

$$\Delta_t^Z(\mathbf{s}) = \beta_{0,t+3}(\mathbf{s}) + \left(\beta_1 + \beta_{1,t+3}(\mathbf{s})\right) W_{t+3}(B) + \epsilon_{t+3}(\mathbf{s}) - \beta_{0,t-5}(\mathbf{s}) - \left(\beta_1 + \beta_{1,t-5}(\mathbf{s})\right) W_{t-5}(B) - \epsilon_{t-5}(\mathbf{s}). \quad (8)$$

In the interest of expediting computation for model fitting, we will simplify (8) so that we regress $\Delta_t^Z(\mathbf{s})$ on the change in Eta-CMAQ. Let $X_t(\mathbf{s})$ denote the current 8-hour average Eta-CMAQ output for each site $\mathbf{s}$ belonging to the grid cell B. Analogous to (6), we define the Eta-CMAQ data differences

$$\Delta_t^x(\mathbf{s}) = 8\left(X_t(\mathbf{s}) - X_{t-5}(\mathbf{s})\right).$$

In fact, for $\mathbf{s} \in B$, $\Delta_t^x(\mathbf{s}) = W_{t+3}(B) - W_{t-5}(B)$.

Figure 2 shows the monitoring data differences for four randomly chosen sites and the Eta-CMAQ data differences for the corresponding grid cells, for one-week period. The plots show good agreement between $\Delta_t^Z(\mathbf{s})$ and $\Delta_t^x(\mathbf{s})$ suggesting that the Eta-CMAQ data differences will be useful predictors of the monitoring data differences. Under the multiplicative forms in (5) and the assumption $\beta_{1,t+3} = \beta_{1,t-5}$, i.e., no temporal effects for the slope process (in fact, we can set the common $\beta_{1,t} = 1$ without loss of generality since it only scales the slope process), we can rewrite the expression (8) in terms of Eta-CMAQ data differences. Hence, our full model is given by

$$\Delta_t^Z(\mathbf{s}) = \beta_{0,t}^*(\mathbf{s}) + \tilde{\beta}_1(\mathbf{s}) \Delta_t^x(\mathbf{s}) + \Delta_t^\epsilon(\mathbf{s}) \quad (9)$$

where $\Delta_t^\epsilon(\mathbf{s}) = \epsilon_{t+3}(\mathbf{s}) - \epsilon_{t-5}(\mathbf{s})$ and

$$\begin{aligned}
\beta_{0,t}^*(\mathbf{s}) &= \beta_0(\mathbf{s})\left(\beta_{0,t+3} - \beta_{0,t-5}\right) = \beta_0(\mathbf{s})\beta_{0,t}^* \\
\tilde{\beta}_1(\mathbf{s}) &= \beta_1 + \beta_1(\mathbf{s}).
\end{aligned} \quad (10)$$

Figure 3 gives a graphical representation for the differencing leading to the proposed model. In the figure we can also see the future Δ's necessary for the current 8-hour average forecasting (prediction of these Δ's is discussed in Section 5). So, first differences enable useful simplification of the downscaler: a space-time process for the intercepts and a purely spatial process for the slopes. With a smaller number of parameters and a straightforward

dependence structure, we reduce the computing time needed for fitting the model and facilitate forecasting the current 8-hour average ozone concentration.

As a special case, we can fix $\beta_0(\mathbf{s}) = 1$ in (10); in fact, this assumption corresponds to modeling the spatio-temporally varying intercept $\beta_{0,t}(\mathbf{s})$ in (1) using an additive form in the spatial and temporal effects. The pure spatial component in the intercept will disappear when we derive the downscaler model for the differences, supporting the assertion above that the additive model in the random effects is less flexible than the multiplicative one.

The spatio-temporally varying intercept $\beta_{0,t}^*(\mathbf{s})$ under the multiplicative assumption in (10) emerges as a zero-mean process with a separable covariance structure which we write as

$$Cov\left[\beta_{0,t}^*(\mathbf{s}), \beta_{0,t'}^*\left(\mathbf{s}'\right)\right] = \sigma^2 \rho^{(s)}\left(\mathbf{s} - \mathbf{s}'; \phi_0\right) \rho^{(t)}\left(t - t'; \varphi\right) \quad (11)$$

where $\rho^{(s)}$ is a valid two-dimensional spatial correlation function and $\rho^{(t)}$ is a valid one-dimensional temporal correlation. Furthermore, the local spatial adjustment $\beta_1(\mathbf{s})$ in (10) is modeled as a zero-mean Gaussian process with covariance structure of the form

$$Cov\left[\beta_1(\mathbf{s}), \beta_1\left(\mathbf{s}'\right)\right] = \xi^2 \rho^{(s)}\left(\mathbf{s} - \mathbf{s}'; \phi_1\right) \quad (12)$$

We acknowledge the simplification associated with the separable specification but the model in (9)-(10) implies that the resulting process for the Δ^Z's does not have a separable covariance function. Indeed, we have

$$Cov\left[\Delta_t^Z(s), \Delta_{t'}^Z\left(s'\right)\right] = \sigma^2 \rho^{(s)}\left(s - s'; \phi_0\right) \rho^{(t)}\left(t - t'; \varphi\right) + + \Delta_t^x(s) \Delta_{t'}^x\left(s'\right) \xi^2 \rho^{(s)}\left(s - s'; \phi_1\right)$$

which is nonstationary. We take $\rho^{(s)}$ in (11) and (12) to be exponential correlation functions, i.e. $\rho^{(s)}(\mathbf{s} - \mathbf{s}'; \phi) = \exp(-\phi \|\mathbf{s} - \mathbf{s}'\|)$ while $\rho^{(t)}$ is the correlation function of an $AR(1)$ model, i.e. $\rho^{(t)}(t - t'; \varphi) = \varphi^{|t-t'|}/(1 - \varphi^2)$.

Lastly, for convenience and, again, to expedite computation, we assume $\beta_{0,t}^*(\mathbf{s})$ and $\beta_1(\mathbf{s})$ independent. Of course, it would be possible to introduce association between intercept and slope using, say the method of coregionalization [12, 13]. However, we do not pursue this further here.

4. Model fitting

It is well-known that it is not possible to consistently estimate the decay and variance parameter in a spatial model with a covariance function belonging to the Matérn family [14]. With exponential covariance functions, the product, $\sigma^2 \phi$ is consistently estimable (identifiable) and spatial interpolation is sensitive to the product $\sigma^2 \phi$ but not to either one individually [15]. For these reasons, along with our ongoing objective of rapid computation for model fitting, we choose optimal values of ϕ and φ offline, using a validation mean square error criterion (see Section 6) and then infer about the variances conditional on these values.

Denote the remaining unknown parameters by $\boldsymbol{\theta} = (\beta_1, \tau^2, \sigma^2, \xi^2)$. For the parameter β_1 we assume a normal prior distribution $N(0, g^2)$ with g^2 taken to be large. For the variance parameters σ^2, ξ^2 and τ^2 we specify independent proper inverse gamma prior distributions $IG(a, b)$; in our implementation we take $a = 2$ and $b = 1$, i.e., a rather vague prior distribution with mean 1 and infinite variance.

4.1. Posterior details

For an observed set of locations $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$ and hours $t = 1, \ldots, (T-3)$, given $\{\beta_{0,t}^*\}$, $\{\beta_0(\mathbf{s}_i)\}$, $\{\beta_1(\mathbf{s}_i)\}$ and $\boldsymbol{\theta}$, the $\Delta_t^Z(\mathbf{s}_i)$ are conditionally independent. Hence, the likelihood is

$$L\left(\theta, \mathbf{B}_0^{(t)}, \mathbf{B}_0^{(s)}, \mathbf{B}_1^{(s)}; \Delta^Z\right) \propto \left(\tau^2\right)^{\frac{(T-3)n}{2}} \exp\left\{-\frac{1}{2\tau^2}\sum_{t=1}^{T-3}\sum_{i=1}^{n}\left(\Delta_t^Z(\mathbf{s}_i) - \beta_{0,t}^*\beta_0(\mathbf{s}_i) - \beta_1\Delta_t^x(\mathbf{s}_i) - \beta_1(\mathbf{s}_i)\Delta_t^x(\mathbf{s}_i)\right)^2\right\}$$

where Δ^Z denotes all the data, $\mathbf{B}_0^{(t)} = \left(\beta_{0,1}^*, \ldots, \beta_{0,T-3}^*\right)'$, $\mathbf{B}_0^{(s)} = (\beta_0(\mathbf{s}_1), \ldots, \beta_0(\mathbf{s}_n))'$ and $\mathbf{B}_1^{(s)} = (\beta_1(\mathbf{s}_1), \ldots, \beta_1(\mathbf{s}_n))'$. The joint posterior distribution is given by

$$\begin{aligned}\pi\left(\theta, \mathbf{B}_0^{(t)}, \mathbf{B}_0^{(s)}, \mathbf{B}_1^{(s)}|\Delta^Z\right) &\propto L\left(\theta, \mathbf{B}_0^{(t)}, \mathbf{B}_0^{(s)}, \mathbf{B}_1^{(s)}; \Delta^Z \times\right. \\ &\quad \times \pi(\beta_1) \times \pi\left(\tau^2\right) \times \pi\left(\sigma^2\right) \times \pi\left(\xi^2\right) \times \\ &\quad \times \pi\left(\mathbf{B}_0^{(t)}\right) \times \pi\left(\mathbf{B}_0^{(s)}\right) \times \pi\left(\mathbf{B}_1^{(s)}\right)\end{aligned}$$

where $\pi(\beta_1)$, $\pi(\tau^2)$, $\pi(\sigma^2)$ and $\pi(\xi^2)$ denote the prior distributions described above. This model is fitted using a Gibbs sampler. The full conditional distributions are developed in the Appendix.

5. Prediction details

Once the model is fitted, we turn to the primary goal of forecasting 8-hour average ozone concentration at the current hour T. According to the definition of $Z_T(\mathbf{s})$ in (3), we will always need to predict three hours into the future in order to forecast current 8-hour average concentration. Equivalently, monitoring data differences are available up to $\Delta_{T-3}^Z(\mathbf{s})$. So, we need to predict $\Delta_{T-2}^Z(\mathbf{s})$, $\Delta_{T-1}^Z(\mathbf{s})$ and $\Delta_T^Z(\mathbf{s})$ in order to forecast $Z_T(\mathbf{s})$, that is,

$$\begin{aligned}Z_T(\mathbf{s}) &= Z_{T-1}(\mathbf{s}) + \Delta_T^Z(\mathbf{s})/8 \\ &= Z_{T-2}(\mathbf{s}) + \Delta_{T-1}^Z(\mathbf{s})/8 + \Delta_T^Z(\mathbf{s})/8 \\ &= Z_{T-3}(\mathbf{s}) + \Delta_{T-2}^Z(\mathbf{s})/8 + \Delta_{T-1}^Z(\mathbf{s})/8 + \Delta_T^Z(\mathbf{s})/8.\end{aligned} \quad (13)$$

Returning to the graphical representation of the model in Figure 3, we see how the available information is used to obtain the forecasts we require. As noted in the Introduction, the Eta-CMAQ forecasts are available 48 hours into the future, so we have the necessary ingredient to make these predictions using the model in (9)-(10). Predictions at new site $\mathbf{s}'$ and hours of interest $T+l$, $(l = -2, -1, 0)$ are based upon the predictive distribution of $\Delta_{T+l}^Z(\mathbf{s}')$. Under our model (9)-(10), $\Delta_{T+l}^Z(\mathbf{s}')$ is conditionally independent of the data Δ^Z up to time T, given $\boldsymbol{\theta}$, $\beta_{0,T+l}^*$, $\beta_0(\mathbf{s})$ and $\beta_1(\mathbf{s}')$ and its distribution is

$$\Delta_{T+l}^Z\left(\mathbf{s}'\right) \sim N\left(\beta_{0,T+l}^*\left(\mathbf{s}'\right) + \tilde{\beta}_1\left(\mathbf{s}'\right)\Delta_{T+l}^x\left(\mathbf{s}'\right), \tau^2\right). \quad (14)$$

Again, the distribution in (14) highlights the contribution of the Eta-CMAQ output $\Delta_{T+l}^x(\mathbf{s}')$ which, as we have noted, is available for these three future hours. The posterior predictive distribution of $\Delta_{T+l}^Z(\mathbf{s}')$ is given by

$$\pi\left(\Delta_{T+l}^{Z}\left(\mathbf{s}'\right)|\Delta^{Z}\right)$$
$$= \int \pi\left(\Delta_{T+l}^{Z}\left(\mathbf{s}'\right)|\beta_{0,T+l}^{*},\beta_{0}\left(\mathbf{s}'\right),\beta_{1},\beta_{1}\left(\mathbf{s}'\right),\tau^{2}\right)$$
$$\times \pi\left(\beta_{0,T+l}^{*}|\mathbf{B}_{0}^{(t)},\varphi\right)$$
$$\times \pi\left(\beta_{0}\left(\mathbf{s}'\right)|\mathbf{B}_{0}^{(s)},\sigma^{2},\phi\right)$$
$$\times \pi\left(\beta_{1}\left(\mathbf{s}'\right)|\mathbf{B}_{1}^{(s)},\xi^{2},\phi\right)$$
$$\times \pi\left(\theta,\mathbf{B}_{0}^{(t)},\mathbf{B}_{0}^{(s)},\mathbf{B}_{1}^{(s)}|\Delta^{Z}\right)d\beta_{0,T+l}^{*}d\beta_{0}\left(\mathbf{s}'\right)d\beta_{1}\left(\mathbf{s}'\right)d\mathbf{B}_{0}^{(t)}d\mathbf{B}_{0}^{(s)}d\mathbf{B}_{1}^{(s)}d\theta.$$

$$(15)$$

The predictive distribution in (15) is sampled by composition. In particular, we need to generate draws for $\beta_{0,T+l}^{*}$, $\beta_{0}(\mathbf{s}')$ and $\beta_{1}(\mathbf{s}')$, conditional on the posterior samples at the observed locations and hours, in order to obtain draws for $\Delta_{T+l}^{Z}\left(\mathbf{s}'\right)$. Given the $AR(1)$ model for $\mathbf{B}_{0}^{(t)}$, we have

$$\beta_{0,T+l}^{*}|\mathbf{B}_{0}^{(t)},\varphi \sim N\left(\varphi\beta_{0,(T+l-1)}^{*},1\right)$$

For the spatially varying intercept, the joint distribution of $\mathbf{B}_{0}^{(s)}$ and $\beta_{0}(\mathbf{s}')$ is a multivariate normal from which the conditional distribution is the univariate normal

$$\beta_{0}\left(\mathbf{s}'\right)|\mathbf{B}_{0}^{(s)},\sigma^{2},\phi \sim N\left(\sum_{i=1}^{n}b_{i}\left(\mathbf{s}'\right)\beta_{0}\left(\mathbf{s}_{i}\right),\sigma^{2}C\left(\mathbf{s}'\right)\right)$$

where

$$b_{i}\left(\mathbf{s}'\right)=\sum_{j=1}^{n}\rho^{(s)}\left(\mathbf{s}'-\mathbf{s}_{j};\phi\right)\left(H^{-1}\left(\phi\right)\right)_{ij}$$

and

$$C\left(\mathbf{s}'\right)=1-\sum_{j=1}^{n}\sum_{i=1}^{n}\rho^{(s)}\left(\mathbf{s}'-\mathbf{s}_{i};\phi\right)\left(H^{-1}\left(\phi\right)\right)_{ij}\rho^{(s)}\left(\mathbf{s}_{j}-\mathbf{s}';\phi\right)$$

Similarly, we generate the random variable $\beta_{1}(\mathbf{s}')$ conditional on the posterior samples at the observed locations. For this, we have

$$\beta_{1}\left(\mathbf{s}'\right)|\mathbf{B}_{1}^{(s)},\xi^{2},\phi \sim N\left(\sum_{i=1}^{n}b_{i}\left(\mathbf{s}'\right)\beta_{1}\left(\mathbf{s}_{i}\right),\xi^{2}C\left(\mathbf{s}'\right)\right)$$

where $b_{i}(\mathbf{s}')$ and $C(\mathbf{s}')$ are defined as above. The conditional means and variances are computationally expensive to compute. However, by fixing the decay parameters ϕ and φ, the quantities b_{i} and $C(\mathbf{s}')$ need only be calculated once and stored; no updating is required in the MCMC, facilitating real-time forecasting.

5.1. Forecast map

Again, our goal is to provide, in a real-time environment, hourly spatial interpolation maps of 8-hour average ozone concentration. To obtain these maps, we need spatial predictions at each Eta-CMAQ grid cell centroid, such as the EPA AIRnow system supplies, roughly 54,000 cells. Given the limited time available to produce plausible predictions at such a large number of grid cell points, formal Bayesian kriging (as in say, Banerjee et al. [16]) will not offer a feasible approach. So, at this last stage, we introduce approximation. Again, this last stage is only for the map making. There will be sufficient time for the foregoing model fitting.

We propose to interpolate the predictions obtained from our model at the n monitoring sites to the Eta-CMAQ grid cell centroids by ordinary kriging, using a fast, available package. In this regard, we can adopt one of the following approaches. The first method is to apply the kriging interpolation both to $Z_{T-3}(\mathbf{s}_i)$ and the posterior predictive samples of $\Delta_{T-2}^Z(\mathbf{s}_i)$, $\Delta_{T-1}^Z(\mathbf{s}_i)$ and $\Delta_T^Z(\mathbf{s}_i)$, with $i = 1, \ldots, n$. Then, the posterior predictive distribution of $Z_T(\mathbf{s})$ at the Eta-CMAQ centroids can be provided by the sum in (13). This approach, however, will be slow and it will introduce large uncertainty to the predictions. Thus, we first sum the last available observation $Z_{T-3}(\mathbf{s}_i)$ and the posterior predictive samples of $\Delta_{T-2}^Z(\mathbf{s}_i)$, $\Delta_{T-1}^Z(\mathbf{s}_i)$ and $\Delta_T^Z(\mathbf{s}_i)$. Then, we obtain the posterior predictive distribution of $Z_T(\mathbf{s})$ at the Eta-CMAQ centroids by kriging. We get the predicted surface of 8-hour average ozone concentration as an average of the posterior predictive distribution of the kriged $Z_T(\mathbf{s})$. A posterior standard deviation map gives a measure of the uncertainty associated with our forecasts.

6. Analysis

We illustrate by modeling data and forecasting using a running window of 24 hours, starting at any given hour. We have investigated longer windows, such as 48-hour and 72-hours. However, the higher computational burden associated to more distant past data is not justified in terms of any improvement in the predictions.

About 5% of values are missing in the monitoring data set. We decided to handle the missingness by removing monitoring sites with at least one missing value in each selected 24-hour window. This choice reflects the structure of the missing values in the data set. As the window changes in time, so do the locations of the missing data. However, in general, missing values occur at monitoring sites for several consecutive hours. This discourages attempts to use 'cheap' imputation; alternatively, a fully model-based imputation would be too computationally expensive. Figure 4 shows the percentage of monitoring sites available to fit the model with respect to 24-hour, 48-hour and 72-hour windows. The 24-hour window enable us to save more than 93% percentage of monitoring sites. So, in addition to being computationally faster, the 24 hour window gains roughly 7% more sites than, say the 48 hour window.

First, we select the decay parameters using the validation criterion described below, recalling that we have set aside data from 70 monitoring stations (Figure 1). For convenience, we set $\phi_0 = \phi_1 = \phi$, imagining that the spatial range for the slope process might agree with that of the intercept process (this simplification is not critical and is really just illustrative). For ϕ and φ, let $\widetilde{\Delta_t^Z}\left(\mathbf{s}_j'\right)$ denote the predicted value at validation site $\mathbf{s}_j'$ for each $j = 1, \ldots, m = 70$ and hours $t = 1, \ldots, (T-3) = 24$.

We employ the Validation Mean Square Error (VMSE)

$$VMSE = \frac{1}{n_v} \sum_{j=1}^{m} \sum_{t=1}^{(T-3)} \left(\Delta_t^Z \left(\mathbf{s}_j' \right) - \widehat{\Delta_t^Z} \left(\mathbf{s}_j' \right) \right)^2 I \left(\Delta_t^Z \left(\mathbf{s}_j' \right) \right) \quad (16)$$

where $n_v = \sum_{j=1}^{m} \sum_{t=1}^{(T-3)} I \left(\Delta_t^Z \left(\mathbf{s}_j' \right) \right)$ is the total number of available observations at the 70 validation sites for the 24 hours. We searched for the optimal value of ϕ among the values, 1.5, 0.5 and 0.25 corresponding to spatial ranges of approximately 185, 560 and 1125 kilometers. For the temporal decay parameter φ, we searched for the optimal value in a grid formed by values of 0.75, 0.85 and 0.95.

For each selected 24-hour window, we compute the VMSE in (16) and we choose the combination of ϕ and φ which leads the smallest VMSE. We experimented with many other values of ϕ and φ learning that the VMSE is not very sensitive to choices close to these optimal values.

We fit the model in (9)-(10) on 24-hour running windows starting at each hour from 8AM to 6PM of August 7th in order to forecast current 8-hour average ozone concentration from 10AM to 8PM on August 8th; this particular temporal window is characterized by a high level of along with variability in ozone concentrations. For each selected window, we predict monitoring data differences at the n available monitoring sites for the three future hours (corresponding to $\Delta_{T-2}^Z (\mathbf{s}_i)$, $\Delta_{T-1}^Z (\mathbf{s}_i)$ and $\Delta_T^Z (\mathbf{s}_i)$) and we forecast the current 8-hour average ozone concentrations ($Z_T(\mathbf{s}_i)$, for $i = 1, \ldots, n$). Then, these forecasts are interpolated to the Eta-CMAQ centroids, as we described in Section 5.1. For example, starting at 8AM of August 7th, we model 24 hourly monitoring data differences from 8AM on August 7th to 7AM on August 8th using data from all available monitoring sites. Predictions of monitoring data differences are computed at the monitoring sites for 8AM, 9AM and 10AM on August 8th and forecasts of the current 8-hour average ozone concentrations at the Eta-CMAQ centroids, associated with 10AM on August 8th, are obtained.

6.1. Results

An example of parameter estimates is shown in Table 1 along with Figures 5 and 6 for the modeling of the data from 10AM of August 7th to 9AM of August 8th.

The significant overall slope β_1 shows the expected positive association between Eta-CMAQ data differences and monitoring data differences. Mean spatial effects $\beta_0(\mathbf{s})$ and $\beta_1(\mathbf{s})$ are shown in Figure 5. Figure 6 shows the 95% credible intervals of the temporal effect. We see the anticipated higher variability for the three hours into the future. The diurnal pattern which characterizes the ozone levels is well reproduced on the first differences scale. Overall, the multiplicative form for $\beta_{0,t}^* (\mathbf{s})$ yields spatio-temporal intercepts that provide an hourly scaling of $\beta_0(\mathbf{s})$. Notably, we observed similar parameters estimates for all other starting hours. We illustrate the current 8-hour average map prediction at 12PM on August 8th in Figure 7 (left panel). The right panel shows the standard deviation map.

As a concluding exercise, we compare the *out-of-sample* predictive performance of the model (9)-(10) and the simpler version obtained by fixing the pure spatial component in the intercept $\beta_0(\mathbf{s}) = 1$. We evaluate Bayesian predictions by computing the mean squared error (MSE), mean absolute error (MAE), empirical coverage and average length of the 95% credible interval on $70 \times 11 = 770$ out-of-sample forecasts. Table 2 reports results for these summary statistics for the two models, revealing little difference except for somewhat

shorter predictive intervals for the *reduced* model. This may be an artifact of the validation sites or may reflect possible overfitting for the *full* model. The empirical coverages agree and are a bit below nominal, suggesting that the intervals are bit short. This is likely be due to the simplifications we make in the model for the differences.

We can offer comparison with AirNow predictions for the same time period. We have to consider this comparison with care for the following reasons. AIRNow makes its forecasts at each monitoring station, treating the stations as independent, building a historical regression at each station, and make a simple local forecast. Then, AIRNow uses a kriging routine to predict to the continental scale. It does not use any computer model output. In particular, for any specified hour, AIRNow uses the subset of monitoring stations that reported for that hour, before kriging; the set of sites employed varies by the hour. So, we can consider two comparisons. Starting with our 717 monitoring stations, holding out 70 of them, leaves us with 647 fitting sites. We make hourly predictions for a subset of these sites as clarified above. So does AIRNow but for a different subset. So, hour by hour, if we consider the intersection of these two subsets, and for the intersection, take our predictions and those of AIRNow, we are able to make a fair, pre-interpolation comparison of forecasts. These results are shown in the first two columns of Table 3 and reveal a roughly 30% improvement in prediction at fitted sites. Interestingly, if we then interpolate hour by hour to the 70 hold-out sites, using a commonly employed kriging R-package '*fields*' (http://www.image.ucar.edu/Software/Fields) we obtain the results in the last two columns of Table 3. We see that the kriging package introduces smoothing such that it reduces the benefit of our modeling approach in terms of interpolated predictive performance. Still we do improve and, in addition we do have a measure of uncertainty through the predictive variance. Indeed, the results from Table 2 show that MSE and MAE for the Bayesian forecast validation at the holdout sites are indistinguishable from the pre-interpolation forecast validation results in Table 3 clarifying the improvement we would expect to see were we able to implement fully model based Bayesian kriging in real-time.

Finally, fitting the faster model of Sahu et al. [1] to our 8-hour average data inputs, we obtained[3] MSE = 75.67 and MAE = 6.90, substantially larger than what we obtained for our models in Table 3.

6.2. Feasibility of real-time computing and discussion

Returning to the motivating objective - the feasibility of our method for real-time use - in terms of offline fitting time and time per hourly update, we note the following. The fitting time is evaluated per iteration of Markov Chain Monte Carlo on an Intel(R) Core(TM)2 Duo CPU E8600 (3.33 GHz, 8 GB RAM). The computing time necessary to fit the model with $\beta_0(\mathbf{s}) = 1$ is about 1.1 seconds per iteration, against 2.2 seconds per iteration for the model with $\beta_0(\mathbf{s}) \quad 1$. So, the simpler version of our model appears more suitable to use within the real-time environment. The hourly update involves the forecasts of current 8-hour ozone concentrations. Typically, only three seconds are required to obtain each posterior predictive sample of $Z_T(\mathbf{s})$ at the Eta-CMAQ centroids, according to our strategy described in Section 5. The code is written in R and we can assert that for the region we have investigated, our approach does work in real-time.

However, it is difficult to say anything more useful regarding moving to national scale where there are about 1400 monitoring sites. We would follow the same path as above but the foregoing run times are not really meaningful; they do not extrapolate. We can say that

---

[3]These summary statistics are based on $50 \times 11 = 550$ forecasts. The corresponding statistics computed over our forecasts for the same hour-site combinations are MSE = 42.31 and MAE= 4.91 for the full model and MSE= 42.08 and MAE= 5.01 for the reduced model.

the proposed model is very linear; the Markov chain Monte Carlo is well-behaved, convergence is rapid. Furthermore, our code is written in R and has not been optimized. We would expect a national version to be written with C+ code, which would run possibly an order of magnitude faster. We also have been using a single machine; a national undertaking would be expected to employ a better hardware environment, at the least, to run on a faster, multi-processor machine. Alternatively, it may prove more attractive to consider regional models and follow, for each region, the same path as we have developed above. In this way we can capture local effects and directly expedite computation by using parallelization.

7. Concluding remarks

We have addressed a specific applied challenge, real-time forecasting of current 8 hour average ozone levels on the scale of the conterminous U.S. We have formulated a downscaler model that works with differences to expedite computation and have shown that it performs very well and appears to be feasible for real-time implementation.

Future work will find us looking at the possibility of introducing real-time temperature data. We will also consider improved 'next day' ozone forecasts. We are also interested in current and next day particulate matter forecasting where new challenges arise because particulate matter is not necessarily collected daily at monitoring sites. A related opportunity will be to provide improved real-time regional forecasts at finer resolution than the national ones, say for urban areas of interest, obtained concurrently with the national forecasts.

8. Disclaimer

The U.S. Environmental Protection Agency's Office of Research and Development partially collaborated in the research described here. Although it has been reviewed by EPA and approved for publication, it does not necessarily reflect the Agency's policies or views.

Acknowledgments

Appendix A. Appendix

The full conditional distributions for the inverse of the variance parameters τ^2, σ^2 and ξ^2 are:

$$
\begin{aligned}
\frac{1}{\tau^2}|rest &\sim Ga\left(a+\frac{(T-3)n}{2}, b+\frac{1}{2}\sum_{t=1}^{T-3}\sum_{i=1}^{n}\left(\Delta_t^Z(\mathbf{s}_i) - \beta_0(\mathbf{s}_i)\beta_{0,t}^* - \beta_1\Delta_t^x(\mathbf{s}_i) - \beta_1(\mathbf{s}_i)\Delta_t^x(\mathbf{s}_i)\right)^2\right) \\
\frac{1}{\sigma^2}|rest &\sim Ga\left(a+\frac{n}{2}, b+\frac{1}{2}\mathbf{B}_0^{'(s)}H^{-1}(\phi)\mathbf{B}_0^{(s)}\right) \\
\frac{1}{\xi^2}|rest &\sim Ga\left(a+\frac{n}{2}, b+\frac{1}{2}\mathbf{B}_1^{'(s)}H^{-1}(\phi)\mathbf{B}_1^{(s)}\right)
\end{aligned}
$$

The full conditional distribution for the global slope parameter β_1 is: $\beta_1|rest \sim N(vg, v)$ where

$$v^{-1} = \frac{1}{\tau^2} \sum_{t=1}^{T-3} \sum_{i=1}^{n} (\Delta_t^x(\mathbf{s}_i))^2 + \frac{1}{g^2}$$

$$g = \frac{1}{\tau^2} \sum_{t=1}^{T-3} \sum_{i=1}^{n} \Delta_t^x(\mathbf{s}_i) \left(\Delta_t^Z(\mathbf{s}_i) - \beta_0(\mathbf{s}_i) \beta_{0,t}^* - \beta_1(\mathbf{s}_i) \Delta_t^x(\mathbf{s}_i) \right)$$

Let $\Delta_t^Z = \left(\Delta_t^Z(\mathbf{s}_1), \ldots, \Delta_t^Z(\mathbf{s}_n) \right)'$ and $\Delta^Z(\mathbf{s}_i) = \left(\Delta_1^Z(\mathbf{s}_i), \ldots, \Delta_{T-3}^Z(\mathbf{s}_i) \right)'$ the vectors that collect spatial series and temporal series, respectively. The full conditional distribution for the intercept spatial effect is a normal distribution $\mathbf{B}_0^{(s)}|rest \sim N(Vd, V)$ where

$$V^{-1} = \frac{1}{\tau^2} \sum_{t=1}^{T-3} \left(\beta_{0,t}^* \right)^2 + \frac{1}{\sigma^2} H^{-1}(\phi)$$

$$d = \frac{1}{\tau^2} \sum_{t=1}^{T-3} \beta_{0,t}^* \left(\Delta_t^Z - \beta_1 \Delta_t^x - D_t^x \mathbf{B}_1^{(s)} \right)$$

and the matrix D_t^x is diagonal with $(D_t^x)_{ii'} = \Delta_t(\mathbf{s}_i)$. For the slope spatial effect we have[4] $\mathbf{B}_1^{(s)}|rest \sim N(Vd, V)$ where

$$V^{-1} = \frac{1}{\tau^2} \sum_{t=1}^{T-3} (D_t^x)^2 + \frac{1}{\xi^2} H^{-1}(\phi)$$

$$d = \frac{1}{\tau^2} \sum_{t=1}^{T-3} D_t^x \left(\Delta_t^Z - \beta_{0,t}^* \mathbf{B}_0^{(s)} - \beta_1 \Delta_t^x \right)$$

Finally, the full conditional distribution for the temporal effect is a normal distribution $\mathbf{B}_1^{(s)}|rest \sim N(Vd, V)$ where

$$V^{-1} = \frac{1}{\tau^2} \sum_{i=1}^{n} (\beta_0(\mathbf{s}_i))^2 + K^{-1}(\varphi)$$

$$d = \frac{1}{\tau_2} \sum_{i=1}^{n} \beta_0(\mathbf{s}_i) \left(\Delta(\mathbf{s}_i)^Z - \beta_1 \Delta(\mathbf{s}_i)^x - \beta_1(\mathbf{s}_i) \Delta(\mathbf{s}_i)^x \right)$$

References

1. Sahu SK, Yip S, Holland DM. A fast bayesian method for updating and forecasting hourly ozone levels. Environmental and Ecological Statistics. 2009; 18:185–207.

2. Sahu SK, Yip S, Holland DM. Improved space-time forecasting of next day ozone concentrations in the eastern us. Atmospheric Environment. 2009; 43:494–501.

3. Kang D, Mathur R, Rao ST, Yu S. Bias adjustment techniques for improving ozone air quality forecasts. Journal of Geophysical Research. 2008; 113 doi:10.1029/2008JD010151.

4. Gelfand, AE.; Sahu, SK. Combining monitoring data and computer model output in assessing environmental exposure. In: O'Hagan, A.; West, M., editors. Handbook of Applied Bayesian Analysis. Oxford University Press; 2010. p. 482-510.

5. Fuentes M, Raftery A. Model evaluation and spatial interpolation by bayesian combination of observations with outputs from numerical models. Biometrics. 2005; 61:36–45. [PubMed: 15737076]

---

[4]We re-use the same symbols for notation simplicity.

6. Poole D, Raftery A. Inference for deterministic simulation models: The bayesian melding approach. Journal of the American Statistical Association. 2000; 95:1244–1255.

7. McMillan N, Holland D, Morara M, Feng J. Combining numerical model output and particulate data using bayesian space-time modeling. Environmetrics. 2010; 21:48–65.

8. Berrocal VJ, Gelfand AE, Holland DM. A spatio-temporal down-scaler for output from numerical models. Journal of Agricultural, Biological and Environmental Statistics. 2010; 14:176–197.

9. Berrocal VJ, Gelfand AE, Holland DM. A bivariate spatio-temporal downscaler under space and time misalignment. Annals of Applied Statistics. 2010; 4:1942–1975. [PubMed: 21853015]

10. Gelfand AE, Kim H-J, Sirmans CF, Banerjee S. Spatial modeling with spatially varying coefficient processes. Journal of the American Statistical Association. 2003; 98:387–396.

11. Berrocal VJ, Gelfand AE, Holland DM. Space-time data fusion under error in computer model output: an application to modeling air quality. Biometrics. 2012; 68:837–848. [PubMed: 22211949]

12. Wackernagel, H. Multivariate Geostatistics: an introduction with applications. 3rd edition. Springer; Berlin: 2003.

13. Gelfand AE, Schmidt AM, Banerjee S, Sirmans C. Nonstationary multivariate process modeling through spatially varying coregionalization. TEST. 2004; 13:263–312.

14. Zhang H. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. Journal of the American Statistical Association. 2004; 99:250–261.

15. Stein, M. Interpolation of Spatial Data: Some Theory for Kriging. Springer Verlag; New York: 1999.

16. Banerjee, S.; Carlin, BP.; Gelfand, AE. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall; CRC: 2004.

Figure 1.
Ozone monitoring sites in the eastern U.S. Dots and crosses represent data and validation sites, respectively.
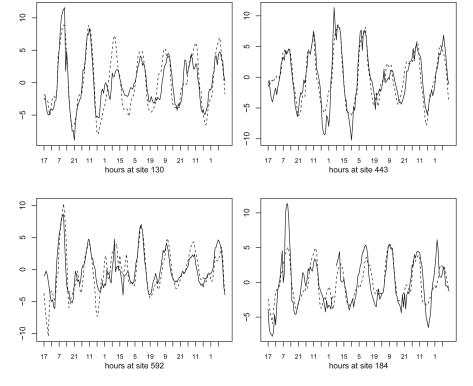
Figure 2.
Monitoring data differences $\Delta_t^Z(\mathbf{s})$ and Eta-CMAQ data differences $\Delta_t^x(\mathbf{s})$ from 4 randomly chosen sites for one-week period.
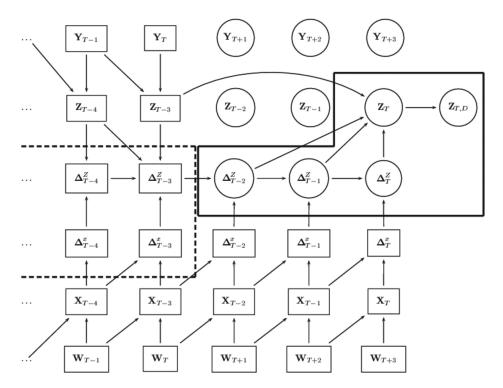
Figure 3.
Graphical representation of our model at the current hour T. □: observed variables. ○: unobserved variables. We model the variables inside the dashed box and we predict the quantities inside the solid box.
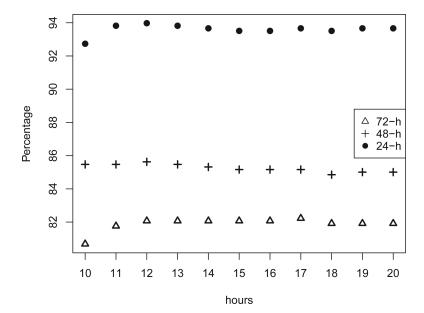
Figure 4.
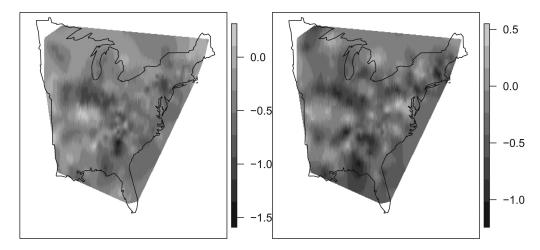Percentage of monitoring sites available to fit the model.

Figure 5.
Mean spatial effects $\beta_0(\mathbf{s})$ (left panel) and $\beta_1(\mathbf{s})$ (right panel) when we model data starting at 10AM of August 7th.
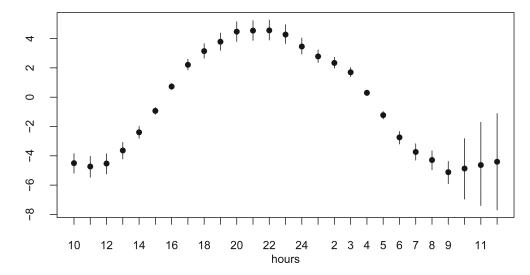
Figure 6.
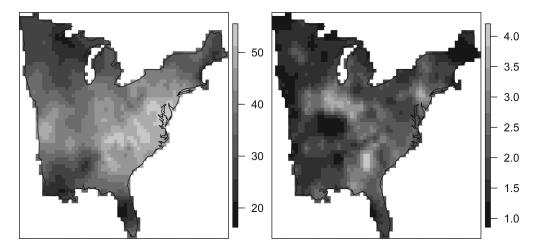95 % credible interval of the temporal component when we model data starting at 10AM of August 7th.

Figure 7.
Current 8-hour average ozone forecast map (left panel) and standard deviation map (right panel) at 12PM of August 8th.

Table 1

Parameter estimates under full model when we model data starting at 10AM of August 7th.

β_1			τ^2			σ^2			ξ^2		
mean	95% interval		mean	95% interval		mean	95% interval		mean	95% interval	
0.634	0.512	0.776	1.470	1.433	1.506	0.499	0.364	0.686	0.453	0.369	0.564

Table 2

Mean square error (MSE), mean absolute error (MAE), empirical coverage and average length of 95% predictive intervals (PI) for full model with $\beta_0(\mathbf{s}) \neq 1$ and reduced model with $\beta_0(\mathbf{s}) = 1$.

	MSE	MAE	Empirical coverage of 95% PI	Average length of 95% PI
$\beta_0(\mathbf{s}) \neq 1$	24.97	3.80	85.7%	15.70
$\beta_0(\mathbf{s}) = 1$	24.66	3.79	85.5%	13.67

Table 3

Mean square error (MSE) and mean absolute error (MAE) for full model, reduced model and AIRNow forecasts.

	Pre-Interpolation		Post-Interpolation	
	MSE	**MAE**	**MSE**	**MAE**
$\beta_0(\mathbf{s}) \quad 1$	25.46	3.91	42.35	4.96
$\beta_0(\mathbf{s}) = 1$	24.43	3.87	41.95	4.95
AIRNow	36.39	4.73	45.72	5.35