Probabilistic Weather Forecasting for Winter Road Maintenance

Veronica J. BERROCAL, Adrian E. RAFTERY, Tilmann GNEITING, and Richard C. STEED

Winter road maintenance is one of the main tasks for the Washington State Department of Transportation. Anti-icing, that is, the preemptive application of chemicals, is often used to keep the roadways free of ice. Given the preventive nature of anti-icing, accurate predictions of road ice are needed. Currently, anti-icing decisions are usually based on deterministic weather forecasts. However, the costs of the two kinds of errors are highly asymmetric because the cost of a road closure due to ice is much greater than that of taking anti-icing measures. As a result, probabilistic forecasts are needed to optimize decision making.

We propose two methods for forecasting the probability of ice formation. Starting with deterministic numerical weather predictions, we model temperature and precipitation using distributions centered around the bias-corrected forecasts. This produces a joint predictive probability distribution of temperature and precipitation, which then yields the probability of ice formation, defined here as the occurrence of precipitation when the temperature is below freezing. The first method assumes that temperatures, as well as precipitation, at different spatial locations are conditionally independent given the numerical weather predictions. The second method models the spatial dependence between forecast errors at different locations. The model parameters are estimated using a Bayesian approach via Markov chain Monte Carlo.

We evaluate both methods by comparing their probabilistic forecasts with observations of ice formation for Interstate Highway 90 in Washington State for the 2003–2004 and 2004–2005 winter seasons. The use of the probabilistic forecasts reduces costs by about 50% when compared to deterministic forecasts. The spatial method improves the reliability of the forecasts, but does not result in further cost reduction when compared to the first method.

KEY WORDS: Cost–loss ratio; Latent Gaussian process; Markov chain Monte Carlo; Numerical weather forecast; Predictive distribution; Spatial dependence.

1. INTRODUCTION

Ice and snow on roads have large impacts. Failure to maintain roads in winter often leads to road closures and hence economic losses (Sherif and Hassan 2004). The Washington State Department of Transportation estimated that between 1992 and 2004 Snoqualmie Pass on Interstate Highway 90 (I-90) was closed 120 hr per year on average, causing an annual loss of at least 17.5 million dollars. Ice and snow also increase the risk of accidents (Norrman, Eriksson, and Lindqvist 2000; Eriksson and Norrman 2001). The crash rate on the I-90 Mountains to Sound Greenway, Washington State's primary east–west bound highway, in the presence of snow is about five times the rate in clear conditions (Federal Highway Administration 2006).

Several strategies can be used for winter road maintenance. Among these, two of the most common are de-icing, in which chemicals are used to melt ice and snow, and anti-icing, a preventive measure that reduces ice by hindering bonds between ice crystals and road pavement. Due to the use of chemicals, both strategies hurt the environment; soil, vegetation, streams, road surface, and vehicles are all damaged by the chemicals used in both de-icing and anti-acing (Shao and Lister 1996; Ramakrishna and Viraraghavan 2005). However, anti-icing is preferred to de-icing on roads with high traffic volume because it reduces total chemical use and allows a higher level of service to the public.

The costs of winter road maintenance are high, but the losses due to road closures are much higher so accurate ice forecasts are needed (Shao 1998; Chapman, Thornes, and Bradley 2001a). Currently, forecasts of road ice come primarily from numerical road prediction models or numerical weather prediction models. Numerical road prediction models take weather forecasts and road condition data as inputs and forecast future road conditions using the surface energy-balance equation, which describes the flux of energy between the atmosphere and the road (Sass 1992). Numerical weather prediction models forecast future weather by integrating coupled differential equations representing the physical processes that govern the atmosphere forward in time. Both kinds of forecasts are deterministic and do not assess uncertainty, which is a critical factor in weather-related decision making (Palmer 2000; Richardson 2000; American Meteorological Society 2002; Gneiting and Raftery 2005; National Research Council of the National Academies 2006; Roulston et al. 2006).

The cost of failing to take anti-icing measures when ice does form on the road is much greater than that of anti-icing when no ice forms. As a result, it will be best to take anti-icing measures when the predicted probability of ice is greater than some threshold that is typically well below 50%. Thus, a good estimate of the probability of ice formation is needed and deterministic forecasts do not provide it. In this article we present

Veronica J. Berrocal is Postdoctoral Fellow, Statistical and Applied Mathematical Sciences Institute (SAMSI), Research Triangle Park, NC 27709 (E-mail: *vjb2@stat.duke.edu*). Adrian E. Raftery is Blumstein–Jordan Professor of Statistics and Sociology, Department of Statistics, University of Washington, Seattle, WA 98195-4322 (E-mail: *raftery@u.washington.edu*). Tilmann Gneiting is Professor of Statistics, Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany. Richard C. Steed is Research Meteorologist, Department of Atmospheric Sciences, University of Washington, Box 351640, Seattle, WA 98195-1640 (E-mail: *steed@atmos.washington.edu*). The authors thank the editor, the associate editor, and three anonymous referees for very helpful discussions. They also thank Cliff Mass and Matthew Stephens for helpful discussions. This research was supported by the U.S. Department of Defense Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research under grant N00014-01-10745, by the National Science Foundation under awards ATM 0724721 and DMS 0706745, and by the Joint Ensemble Forecasting System (JEFS) under subcontract S06-47225 from the University Corporation for Atmospheric Research.

Journal of the American Statistical Association
June 2010, Vol. 105, No. 490, Applications and Case Studies
DOI: 10.1198/jasa.2009.ap07184

methods for estimating this, building on the work of Mass et al. (2003), who documented the joint effort of the Department of Atmospheric Sciences at the University of Washington and of the Washington State Department of Transportation to monitor and provide forecasts of road ice for I-90 by postprocessing numerical weather forecasts.

The physical process that produces ice on roads is complex and involves the interaction of weather with road surface conditions (Chapman, Thornes, and Bradley 2001a; Thornes, Cavan, and Chapman 2005). Here we simplify it and we assume that ice will form at a point on the road if precipitation occurs and the air temperature is at or below freezing. Starting from numerical weather forecasts, we develop a model for the joint predictive probability distribution of temperature and precipitation, which yields the probability of ice formation.

The problem of winter road maintenance is intrinsically spatial. We examine whether spatial dependence in the forecast errors ought to be modeled explicitly by comparing two models, one that does not take account of spatial correlation and one that does. The latter produces joint predictive probability distributions of future temperature and precipitation over space that account for the spatial correlation of the forecast errors.

In Section 2, we describe the data and introduce the statistical models and the estimation methods that we use. In Section 3, we give verification results and in Section 4 we discuss other approaches that were developed for this problem and possible limitations of our methodology.

2. DATA AND METHODS

2.1 Road Maintenance Problem

The I-90 Mountains to Sound Greenway is one of the main arteries of the State of Washington, with a traffic volume of 20 million vehicles per year. The highway crosses the Cascade Mountains with large changes in altitude and connects the urban centers of the Puget Sound with farmlands in eastern Washington; it is an important route for the regional economy. During the winter months, from October to March, I-90 is often congested because of poor driving conditions due to snow accumulation and ice. To provide a safe transit to travelers, the Washington State Department of Transportation employs preventive anti-icing measures during the winter season. In this study, we use data from 10 meteorological stations along a 140 km section that includes Snoqualmie Pass. Figure 1 shows the altitude profile for this section of I-90.

The data span the three winter seasons of 2002–2003, 2003–2004, and 2004–2005 and include minimum temperature and precipitation during the 3-hr interval 1:00 a.m. through 4:00 a.m. Hereafter, we use the term “temperature” to refer to minimum temperature between 1:00 a.m. and 4:00 a.m. Since this is the coldest time interval of the day, it is very likely that, should there be road ice, it will occur during this interval. Precipitation occurrence was recorded if there was at least 0.01 inch of precipitation.

The data comprise 438 days of observations: 131 during the 2002–2003 winter season, 159 during the following winter season, and 148 during the 2004–2005 winter season. The number of observations per station ranges from a minimum of 169 to a maximum of 363, with an average of 307 and a median of 323. There was no discernible pattern to the missing data and our methods do not require us to impute values for the missing observations. Very few observations were available during the month of October. As a consequence, in evaluating the predictive performance of our method, we compare predictions for the months of November through March with the observations.

As the basis for our forecast, we use 12-hr-ahead model runs of temperature and precipitation produced by the fifth-generation Pennsylvania State University/National Center for Atmospheric Research Mesoscale Model (MM5) (Grell, Dudhia, and Stauffer 1994). The model was run by the University of Washington Department of Atmospheric Sciences with initial and boundary conditions supplied by the U.K. Meteorological Office (Grimit and Mass 2002; Eckel and Mass 2005). The forecasts were generated on a 12 km grid and then bilinearly interpolated to the observation sites. Bilinear interpolation is a standard technique for downscaling numerical model output to the point level and it is considered an integral part of the numerical weather prediction model. It performs better than naive interpolation (Shao, Stein, and Ching 2007; Jun, Knutti, and Nychka 2008) and it does not involve using observational data, just the numerical model output at the grid level. The numerical forecasts that we use are based on the information available at 4:00 p.m. the day before, which is typically the most recent available to managers deciding whether to take anti-icing measures overnight.

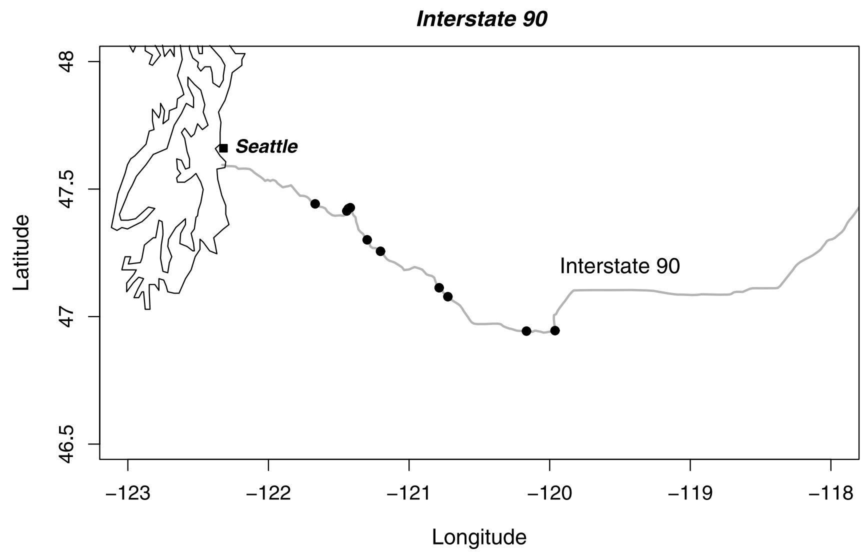
Figure 2 shows weather data at the Alpental station. This is close to Snoqualmie Pass and is one of the highest stations along I-90. The MM5 model predicted temperature quite accurately, with a mean absolute error of the temperature forecast at Alpental during the winter season 2002–2003 of 1.6°C. The precipitation forecasts at Alpental also performed well during the 2002–2003 winter season. For example, in 59 of the 64 cases in which more than 0.01 inch of rain was predicted, precipitation was indeed observed. In 33 of the 50 cases in which the forecast was for no precipitation, the forecast was correct and there was no precipitation.

Most numerical road models, such as ICEBREAKER (Shao and Lister 1996), assume that ice forms on the road if the road surface temperature falls below freezing and the road surface is wet. Reliable direct observations of the conditions of the road pavement for the section of I-90 considered in this study are not available and so we use air temperature and precipitation as proxies for road temperature and road wetness. Adapting the definition of road ice used in numerical road models to the proxy variables, we say that a point along I-90 experiences ice formation if the temperature is equal to or less than 0°C and there is precipitation. Similarly, we say that a section of I-90 experiences ice formation if there is at least one point in the section at which the temperature is at most 0°C and there is precipitation. Our goal is to produce probabilistic forecasts of ice formation along a section of I-90.

2.2 Statistical Model

We now describe our statistical model for ice formation at a given time. The time is fixed and so is not explicitly included in our notation. We write $Y(s)$ for the observed temperature at a site s along the portion $\mathcal{I}$ of I-90 shown in Figure 1. We let

(a)



(b)

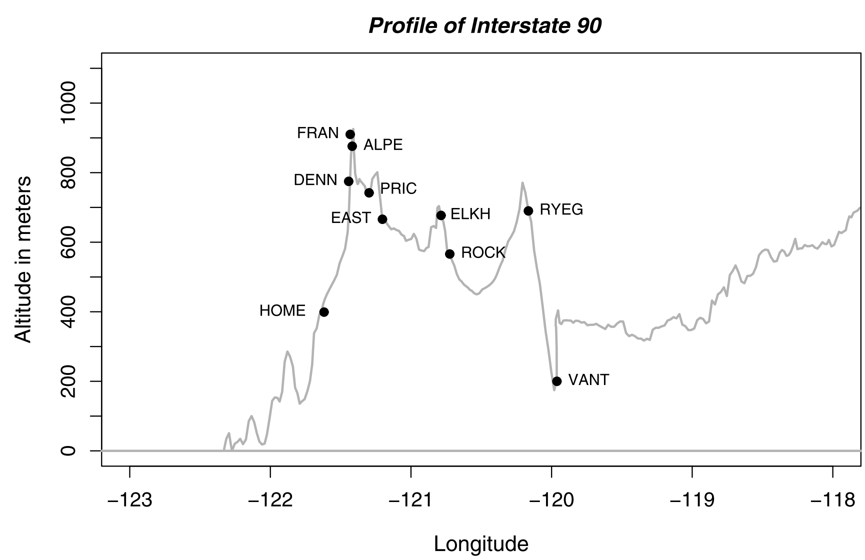


Figure 1. (a) Meteorological stations along I-90. (b) Altitude profile along I-90.

$W(s) = 1$ if precipitation occurs at s and 0 if not, and we say that ice forms on the road at s if

$$Y(s) \leq 0^\circ\text{C} \qquad \text{and} \qquad W(s) = 1. \tag{1}$$

Our goal is to produce probabilistic forecasts of ice formation simultaneously at all sites s on a grid of points along $\mathcal{I}$. We write $\tilde{Y}(s)$ for the 12-hr-ahead forecast of temperature at s and we follow Sloughter et al. (2007) in using the cube root of the forecast of accumulated precipitation, which we denote by $\tilde{W}(s)$, as a predictor.

Given our definition of ice formation, in order to produce probabilistic forecasts of ice, we need to specify a joint model for temperature and precipitation occurrence. We assessed the dependence between temperature and precipitation occurrence given the forecasts by dividing the observations into groups with similar values of the forecasts. Within each group we find no evidence of dependence. We, therefore, assume that temperature and precipitation occurrence are conditionally independent of one another given the forecasts. We now specify two models, each of them for the joint distribution of temperature and precipitation occurrence. The two models differ in that one models the spatial correlation in the forecast errors and the other one does not. Given the spatial nature of the data, we are interested in finding out whether modeling the spatial correlation improves predictive performance.

In our first model, which we call the marginal model, we assume that temperatures at different locations are independent of one another, as are occurrences or not of precipitation, given the forecasts of temperature and accumulated precipitation. To remove systematic nonstationary variation, we include a regression-based adjustment of the mean temperature field on latitude, longitude, and elevation, similarly to Handcock and Wallis (1994) in their analysis of the average U.S. winter tem-

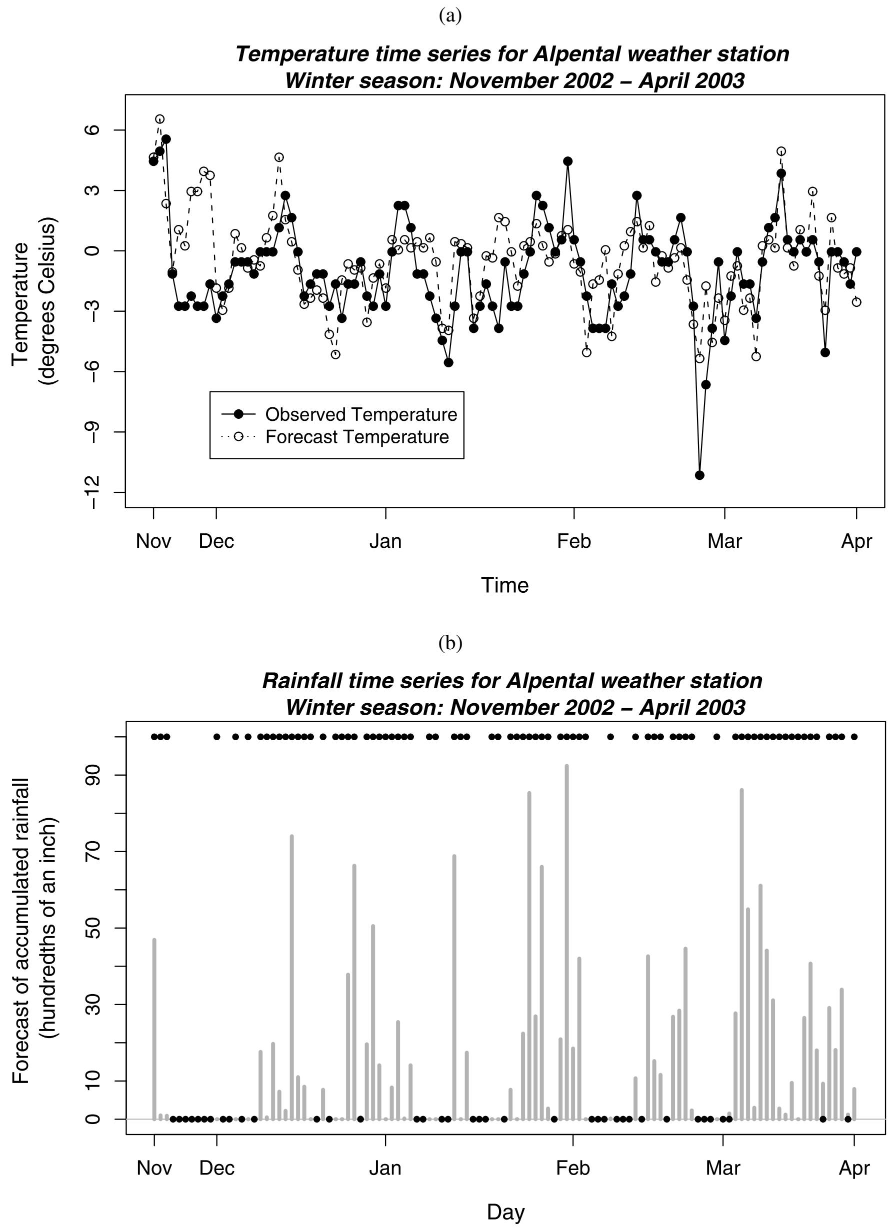


Figure 2. (a) Temperature at the Alpental station during the 2002–2003 winter season. The solid line shows the observed temperature, the dashed line the forecast. (b) Precipitation occurrence (black dots) and forecasts of accumulated precipitation (grey bars) at Alpental during the 3-hr interval 1:00 a.m. to 4:00 a.m.

perature. The marginal model for temperature is then

$$Y(s) = \gamma_0 + \gamma_1 \tilde{Y}(s) + \gamma_2 \,\mathrm{lat}(s) + \gamma_3 \,\mathrm{lon}(s) + \gamma_4 \,\mathrm{ht}(s) + \epsilon(s), \tag{2}$$

where lat, lon, and ht denote the centered latitude and longitude (in degrees) and elevation (in meters), γ_0, γ_1, γ_2, γ_3, and γ_4 are regression coefficients, and $\epsilon(s)$ is a mean-zero Gaussian process with

$$\mathrm{Cov}(\epsilon(s), \epsilon(t)) = \varsigma^2 \delta_{st}, \tag{3}$$

where $\delta_{st} = 1$ if $s = t$ and 0 otherwise.

We now describe our spatial model, which takes account of spatial correlation in forecast errors. The mean temperature field is modeled using the same regression adjustment as in the marginal model. However, the temperature error field is assumed to have a stationary and isotropic exponential covariance structure with nugget, that is,

$$Y(s) = \alpha_0 + \alpha_1 \tilde{Y}(s) + \alpha_2 \,\mathrm{lat}(s) + \alpha_3 \,\mathrm{lon}(s) + \alpha_4 \,\mathrm{ht}(s) + \xi_T(s), \tag{4}$$

where $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, and α_4 are regression coefficients and $\xi_T(s)$ is a mean-zero stationary and isotropic Gaussian process with covariance function

$$\mathrm{Cov}(\xi_T(s), \xi_T(t)) = \rho^2 \delta_{st} + \tau^2 \exp\left(-\frac{|s-t|}{r}\right), \tag{5}$$

where $|s-t|$ is the Euclidean distance between the two sites and δ_{st} is defined earlier. The covariance parameters are the nugget effect, ρ^2, the sill, $\sigma^2 = \rho^2 + \tau^2$, equal to the marginal variance, and the range, r, which specifies the rate at which the exponential correlation decays (Cressie 1993; Chilès and Delfiner 1999).

We now describe our models for precipitation occurrence. In the marginal model we assume conditional spatial independence for precipitation occurrence given the forecast of precipitation, so that

$$W(s) = \begin{cases} 1 & \text{with probability } \pi(s), \\ 0 & \text{with probability } 1 - \pi(s), \end{cases} \tag{6}$$

where

$$\log\left(\frac{\pi(s)}{1 - \pi(s)}\right) = \lambda_0 + \lambda_1 \tilde{W}(s) + \lambda_2 \,\mathrm{ht}(s), \tag{7}$$

and $W(s)$ and $W(t)$ are independent given $\tilde{W}(s)$, $\tilde{W}(t)$, $\mathrm{ht}(s)$, and $\mathrm{ht}(t)$. The inclusion of latitude and longitude made no difference in predictive performance so these variables were not included. Essentially, the marginal model for precipitation occurrence is a logistic regression model with the cube root of the forecast of accumulated precipitation and elevation as predictor variables.

In specifying our spatial model for precipitation occurrence we follow Albert and Chib (1993) and extend their hierarchical model for independent binary data to spatially dependent binary data. We postulate a latent Gaussian process $Z(s)$ that regulates precipitation occurrence. If the latent variable $Z(s)$ is greater than 0, then there is precipitation at the site; otherwise there is no precipitation.

This is basically equivalent to the Tobit model (Tobin 1958), which was adapted to precipitation modeling by Bardossy and Plate (1992) and subsequently used by Hutchinson (1995), Sansò and Guenni (1999, 2000, 2004), Allcroft and Glasbey (2003), and Rappold, Gelfand, and Holland (2008). Unlike in the model of Bardossy and Plate (1992), where the latent Gaussian random field is conveniently power-transformed and truncated to reproduce the long right tail of the rainfall distribution, in our spatial model the latent Gaussian process is not transformed. This is because our model deals only with the binary random variable of precipitation occurrence and not with the rainfall amount.

The mean of the latent Gaussian process $Z(s)$ is a linear combination of the cube root of the forecast amount, and elevation. The covariance structure of $Z(s)$ is modeled using an exponential covariance function. Other covariance structures can also be used. To ensure identifiability, the marginal variance of $Z(s)$ is set equal to 1 and the only covariance parameter is the range parameter θ, which gives the rate of decay of the spatial correlation.

The complete spatial model for precipitation occurrence is thus:

$$W(s) = \begin{cases} 1 & \text{if } Z(s) > 0, \\ 0 & \text{otherwise}, \end{cases} \tag{8}$$

$$Z(s) = \beta_0 + \beta_1 \tilde{W}(s) + \beta_2 \,\mathrm{ht}(s) + \xi_P(s), \tag{9}$$

where $\xi_P(s)$ is a mean-zero unit variance Gaussian process with covariance function

$$\mathrm{Cov}(\xi_P(s), \xi_P(t)) = \exp\left(-\frac{|s-t|}{\theta}\right). \tag{10}$$

Both our spatial models, for temperature and precipitation, account for the temporal and spatial dependence in the data. The spatial dependence is modeled explicitly in the residual processes $\xi_T(s)$ and $\xi_P(s)$, while the temporal dependence is accounted for by the mean functions of the Gaussian processes $Y(s)$ and $Z(s)$ in Equations (4) and (9), respectively. Temperature and precipitation are temporally autocorrelated, but the forecasts largely account for this and the forecast errors have essentially no autocorrelation. We discuss this further in Section 4.

2.3 Model Fitting

For each day, the parameters of the statistical models in Equations (2) through (10) were estimated using data from a “sliding window” training period consisting of the previous N days.

The parameters of the marginal model in Equations (2) and (3) were estimated by fitting a linear regression of observed temperature on latitude, longitude, elevation, and the forecast temperature. The parameters of the marginal model in Equations (6) and (7) were estimated by fitting a logistic regression of the observed precipitation occurrence on the cube root of the forecast accumulated precipitation and elevation. The data used to fit both regressions were observations and forecasts for the 10 stations along $\mathcal{I}$ from the previous N days. Our choice of the length N of the training period is explained in Section 2.4.

For the spatial model, as synchronous Markov chain Monte Carlo schemes often behave poorly due to weak identifiability and extremely slow mixing (Sahu, Gelfand, and Holland 2006, p. 70), we proceeded in two stages: first, the regression coefficients $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, and α_4 and the marginal variance σ^2 were fitted, then the covariance parameters ρ^2, τ^2, and r were estimated.

Since the training period is a sliding window consisting of the previous N days, only a limited amount of data are available. This can make the ordinary least-squares estimates of the regression coefficients and of the marginal variance unstable, as shown in Figure 3. Therefore, we used a Bayesian framework, adopting the posterior means as estimates of the regression coefficients, $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, and α_4, and of the marginal variance, σ^2. Specifically, our estimates are based on the following Bayesian model:

$$\begin{aligned} Y(s) &= \alpha_0 + \alpha_1 \tilde{Y}(s) + \alpha_2 \,\mathrm{lat}(s) \\ &\quad + \alpha_3 \,\mathrm{lon}(s) + \alpha_4 \,\mathrm{ht}(s) + \xi_T'(s), \\ \alpha &= (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)^T \mid \sigma^2 \sim \mathrm{MVN}_5(\boldsymbol{\eta}, \sigma^2 \boldsymbol{\Omega}), \\ \sigma^2 &\sim \mathrm{InvGamma}(\nu, \psi), \end{aligned} \tag{11}$$

where $\xi_T'(s)$ is a mean-zero Gaussian process with covariance function $\mathrm{Cov}(\xi_T'(s), \xi_T'(t)) = \sigma^2 \delta_{st}$, $\boldsymbol{\eta}$ is a vector with components $(\eta_0, \ldots, \eta_4)$, and $\boldsymbol{\Omega}$ is a diagonal matrix with diagonal elements $(\omega_0^2, \ldots, \omega_4^2)$. This is a simplification in that the spatial structure is not included in this model; however, the verification results in Section 3 indicate that this does not hurt the predictive performance.

We based the prior distributions of $\alpha_0, \alpha_1, \ldots, \alpha_4$ and σ^2 on the data from the first winter season available to us, that of 2002–2003. For each day in the 2002–2003 winter season, we estimated the regression parameters by fitting a linear regression with data from the previous N days. The prior mean and

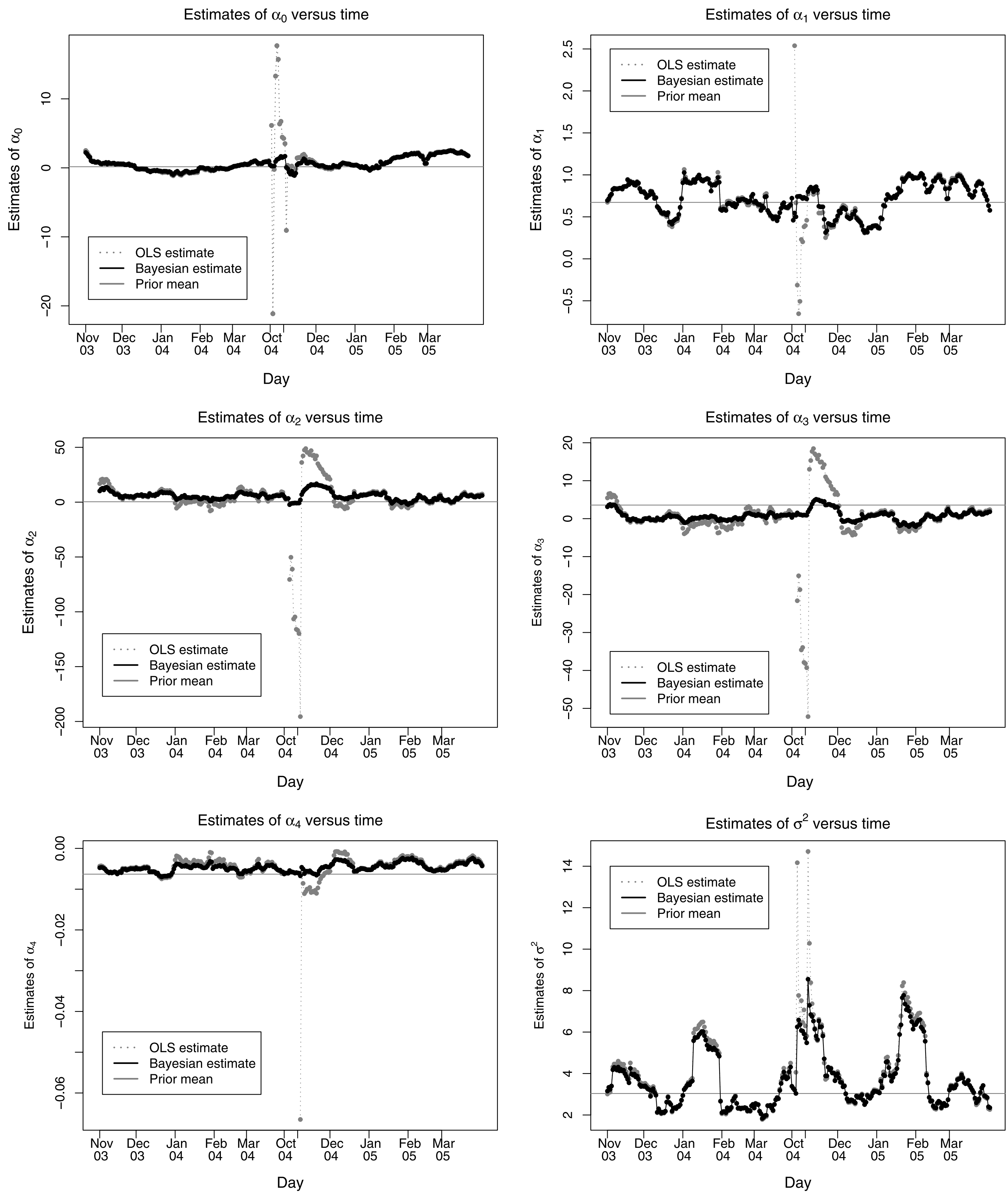


Figure 3. Bayesian estimates of $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4$, and σ^2 versus time (solid black line). The ordinary least-squares estimate (dashed grey line) and the prior mean (solid grey line) are also shown.

prior standard deviation of α_0 were the mean and twice the standard deviation of the resulting estimates of α_0, and similarly for $\alpha_1, \ldots, \alpha_4$. Table 1 reports these prior values for a training period of length $N = 20$.

To specify the hyperparameters of the inverse Gamma prior distribution of σ^2, we minimized the sum of squared deviations between the cumulative distribution function of an inverse Gamma distribution and the empirical distribution of the esti-

Table 1. Prior mean and prior standard deviation of α_0, α_1, α_2, α_3, and α_4 for training period length $N = 20$

	α_0	α_1	α_2	α_3	α_4
Prior mean	0.164	0.672	0.366	3.578	−0.006
Prior standard deviation	0.701	0.205	0.473	1.166	0.001

mates of σ^2. The resulting 10th, 50th, and 90th percentiles of the prior distribution of σ were 1.36°C, 1.67°C, and 2.12°C. The posterior means of the regression parameters are given by standard closed form expressions (e.g., Gelman et al. 2004, chapter 14). As can be seen from Figure 3, the use of a Bayesian approach stabilized the ordinary least-squares estimates when they were unstable, but otherwise hardly changed them at all, which was our goal.

To estimate the covariance parameters ρ^2 and r, which were assumed to be constant in time, we used the data from the 2002–2003 winter season. We constructed the empirical variogram of the residuals of the linear regression of the observed temperature on the centered latitude, longitude, and elevation, and on the forecast of temperature. We then fitted a parametric exponential variogram with nugget effect to the empirical variogram using weighted least-squares (Cressie 1993), where the weights were the numbers of times each pair of stations had been observed simultaneously. Figure 4 shows the empirical variogram of the temperature residuals and the fitted exponential variogram model.

The marginal variance σ^2 varies with time and was re-estimated daily. We ascribe the variability over time of the marginal variance to the variability of the variance of the continuous component in Equation (5). We thus model ρ^2, the small scale variability, as constant in time, while we allow τ^2 to change with time and re-estimate it daily. From Equation (5) it follows that, for each day, we can obtain a new estimate of τ^2 by subtracting the estimate of ρ^2 from the estimate of σ^2.

In estimating the parameters of the spatial model for precipitation occurrence, the parameters in Equations (8), (9), and (10) are estimated for each day in a Bayesian way using data from a training period consisting of the previous N days. We use the following priors for β and θ:

$$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T \sim \mathrm{MVN}_3(\boldsymbol{\mu}, \mathbf{V}), \tag{12}$$

$$\theta \sim \mathrm{Unif}(0, 1000). \tag{13}$$

As we did for temperature, we use spread-out prior distributions with hyperparameters based on data from the 2002–2003 winter season. For each day in the 2002–2003 season, we fit a logistic regression of the observed precipitation occurrence on the cube root of the forecast of accumulated precipitation and on elevation to data from a training period made up of the previous N days. This yielded a set of estimates of β_0, β_1, and β_2. The prior mean $\boldsymbol{\mu}$ then consisted of the means of the estimates of β_0, β_1, and β_2, while $\mathbf{V}$ was diagonal with diagonal elements equal to four times the empirical variances of the estimates.

For each day, the parameters of the spatial hierarchical model for precipitation, given by Equations (8), (9), (10), (12), and (13), were estimated using a Markov chain Monte Carlo algorithm (Gelfand and Smith 1990). Since a closed form for the full conditional distributions of $Z(s)$ and $\boldsymbol{\beta}$ is available, we used a Metropolis–Hastings step to update the covariance parameter θ and a Gibbs sampler algorithm to update all other parameters.

2.4 Choice of Training Period

In choosing the length of the training period there is a trade-off: a shorter period allows changes in the atmosphere to be taken into account more promptly, but on the other hand it reduces the amount of data available for estimation of the parameters.

To choose the length of the training period we used the root mean square error (RMSE) of the temperature predictions and the Brier score (Brier 1950) for the predictions of probability of precipitation, both obtained using training periods of lengths $N = 5, 10, \ldots, 60$ days. In both cases the predictions were made

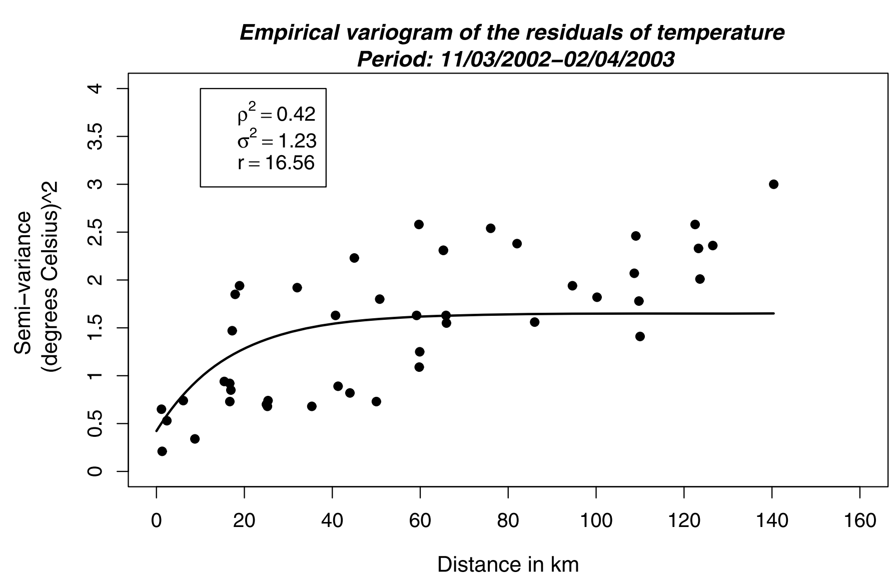


Figure 4. Empirical variogram of the temperature residuals with fitted exponential variogram model. Each dot represents a pair of stations.

at individual sites. The Brier score is defined as:

$$\text{Brier score} = \frac{1}{M}\sum_{i=1}^{M}(o_i - f_i)^2, \tag{14}$$

where M is the total number of predictions, o_i is the ith observed event (1 if the event occurred, 0 otherwise) and f_i is the forecast probability that the ith event will occur. It is negatively oriented, that is, lower is better.

Figure 5 shows the RMSE and the Brier score of the predictions as a function of the training period length. Both plots seems to indicate that 20 days is a good length for the training period: the magnitude of the RMSE error decreased noticeably as the length N of the training period increased up to 20 days, while the Brier score attained its minimum at 20 days. Beyond 20 days, there was not much gain in using a longer training period and the quality of the predictions of temperature worsened slightly for training periods longer than 30 days.

We, therefore, used a training period of 20 days to estimate the model parameters for both temperature and precipitation. Training periods of 25 to 30 days were also found to be adequate for the postprocessing of numerical forecasts of temperature and precipitation at short lead times by Raftery et al. (2005), Sloughter et al. (2007), Hagedorn, Hamill, and Whitaker (2008), and Hamill, Hagedorn, and Whitaker (2008).

2.5 Generating Forecasts

To produce probabilistic forecasts of ice formation along the section $\mathcal{I}$ of I-90 shown in Figure 1, we discretized and simulated from the joint predictive distribution of temperature and precipitation occurrence at 104 points along the highway. The distances between neighboring points range from 1.3 km to 2.0 km and cover a section of I-90 that is about 140 km long.

Realizations of the temperature and precipitation occurrence fields were obtained by simulating from the corresponding stochastic processes: the temperature [Equations (2) and (3)],

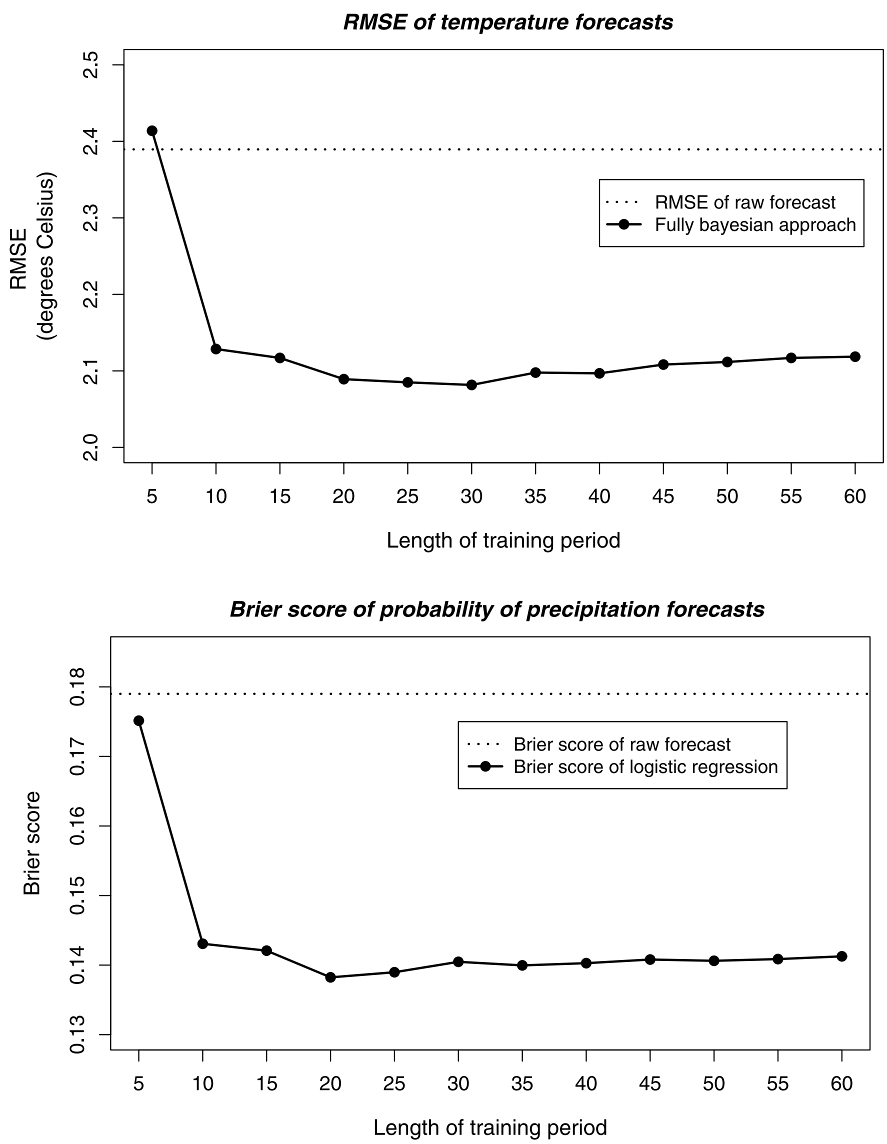


Figure 5. RMSE of temperature forecasts and Brier score for predictions of the probability of precipitation versus training period length, for the 2003–2004 and 2004–2005 winter seasons.

and precipitation occurrence [Equations (6) and (7)] specifications for the marginal model, and the temperature [Equations (4) and (5)] and precipitation occurrence [Equations (8), (9), and (10)] specifications for the spatial model. In both cases, we used parameters estimated from the training data for the simulations. We defined a realization as forecasting ice at one of the 104 points on the road if it had both freezing temperature and nonzero precipitation at that point. The probability of ice on the road at that location was then the proportion of realizations forecasting ice.

Figure 6 maps these probabilities for one day for both models. Not surprisingly, they are similar between the two models because the real differences between the models are for the joint distribution of ice at different places, not for the probability of ice at one place.

To verify the probabilistic forecasts, we used the observations at the 10 stations, for which the forecast probabilities can be computed analytically. We did not observe directly whether there was ice on the road, and instead, we say that ice occurred if the temperature was at or below freezing and there was precipitation.

Comparisons of forecasts and observations at individual sites assess the performance of the predictive distributions only marginally. To evaluate the probabilistic forecasts of the ice field as a whole, we forecast ice formation simultaneously at all 10 observation sites. This was done using the procedure described at the beginning of this section, replacing the 104 points with the 10 observation sites. For a given realization, we say that ice was predicted along the section $\mathcal{I}$ of I-90 if ice formation was forecast at at least 1 of the 10 sites. We say that ice was observed along $\mathcal{I}$ if temperature was at or below freezing and there was precipitation at at least one of the stations.

To distinguish between the two different types of verification—marginal and spatial—we use the expression "ice at observation sites" for verification results relative to forecasts of ice at the individual sites, and "ice along $\mathcal{I}$" to refer to forecasts of ice simultaneously at the 10 observation sites.

3. RESULTS

We now compare the out-of-sample predictive performance of our probabilistic forecasting methods for the 2003–2004 and 2004–2005 winter seasons, for individual sites and for the entire section $\mathcal{I}$ of I-90, with that of several other methods. One of them is the deterministic forecast from the numerical weather prediction model, which we call the "raw" forecast. We define this as predicting ice if it forecasts both freezing temperature and nonzero precipitation.

To determine the impact of a simple bias correction on the predictive performance of the raw forecast, we compare the predictive performance of our methods to that of the bias-corrected raw forecast. We refer to this forecast as the bias-corrected forecast. The bias correction was performed on a daily basis and it was carried out using simple linear regression and data from a 20-day "sliding window" training period.

We also compare our probabilistic forecasts to those from an ensemble of numerical forecasts, the University of Washington Mesoscale Ensemble (UWME; Eckel and Mass 2005). This is an eight-member ensemble generated in the Department of Atmospheric Sciences at the University of Washington, which is obtained by running the MM5 numerical weather prediction model with eight distinct sets of initial and boundary conditions, supplied by major domestic and foreign weather agencies on the basis of past and current observations. The uncertainty captured by the ensemble thus represents the uncertainty in initial and boundary conditions. We also compare our methods to a bias-corrected version of the UWME in which each member of the ensemble was adjusted for bias, using simple linear regression and a 20-day training period.

Our main measure of performance is the Brier score defined in Section 2.4. The Brier score can be decomposed into uncertainty, reliability, and resolution components and equals uncertainty plus reliability minus resolution (Murphy 1973). Specif-

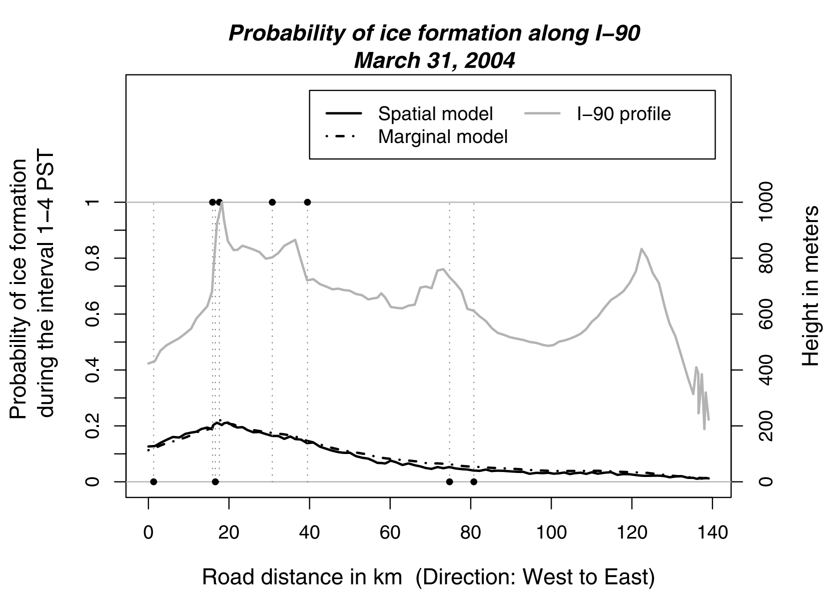


Figure 6. Probability of ice formation along I-90 on March 31, 2004, as forecast by the spatial model (solid black line) and the marginal model (dot–dash black line). The profile of I-90 is shown in grey, while observations are represented by black dots.

ically, if the forecast probabilities take values $p_1, \ldots, p_K$ then

$$\text{Brier score} = \underbrace{\bar{o}(1-\bar{o})}_{\text{uncertainty}} + \underbrace{\frac{1}{M}\sum_{k=1}^{K} n_k (p_k - \bar{o}_k)^2}_{\text{reliability}} - \underbrace{\frac{1}{M}\sum_{k=1}^{K} n_k (\bar{o}_k - \bar{o})^2}_{\text{resolution}},$$

where n_k is the number of times that p_k is forecast, $\bar{o}_k$ is the respective observed relative event frequency, and $\bar{o}$ is the overall relative event frequency. In cases like ours, in which the probability forecast is a continuous variable, the decomposition depends on a binning of the forecast values and is approximate only, but typically is very close to being exact. The uncertainty component measures the inherent uncertainty in the observations and is independent of the forecasts. The reliability component measures the deviation of the reliability curve from the diagonal. It addresses calibration, that is, the statistical consistency between the forecasts and the observations, and is negatively oriented. The resolution component measures the ability of the forecast to distinguish between prior situations that will lead to the occurrence or nonoccurrence of the event and is positively oriented.

Table 2 shows that for the probability that ice forms at individual locations, both of our probabilistic forecasting methods substantially outperformed the raw and the ensemble forecasts. Figures 7 and 8 show reliability diagrams for the probabilistic forecasts of ice formation provided by the marginal model, the spatial model, the UWME, and the bias-corrected UWME. Our methods were superior to both ensembles, which were very underdispersed and particularly unreliable for high predicted probabilities. At individual locations along I-90, the marginal and spatial methods performed similarly, as expected. However, for probability forecasts of the spatial aggregate "ice formation along $\mathcal{I}$," the spatial model was more reliable: the marginal model tended to overestimate the probability of ice formation.

We now compare the forecasting methods in economic terms. In the case of the probability forecasts, we assume that anti-icing measures, costing C, are taken whenever the probability of road ice is greater than a given threshold. A loss L is incurred when no anti-icing measures are taken, but ice does form on the roadway. Various estimates of the economic loss L associated with the closure of I-90 have been reported, ranging from 1 to 18 million dollars per day (Paulson 2001). The cost C of anti-icing measures was estimated at about 100,000 dollars per day.

The threshold probability used to decide whether or not to take anti-icing measures equals the cost–loss ratio, $R = C/L$. Table 3 shows contingency tables cross-classifying action (anti-icing measures or not) against outcome (road ice or not) for each of seven different forecasting methods when $R = 0.1$. We have discussed six of these forecasts: the raw forecast, the bias-corrected forecast, the two ensembles (UWME and bias-corrected UWME), and our two methods. The seventh forecast we consider, the naive forecast, is the same in every instance. If the relative frequency of road ice over a prior winter season is above the threshold probability R, then the naive forecast always predicts road ice; otherwise it always predicts no road ice. The marginal probability of ice formation over the 2002–2003 winter season was 0.07 (9 instances out of 131 forecast events) and so the naive forecast always predicts ice if $R < 0.07$ and never predicts ice if $R \geq 0.07$.

The contingency tables enable us to determine what the total cost associated with winter road maintenance decisions will be if we had used these forecasts during the two winter seasons. We write n_{00}, n_{01}, n_{10}, and n_{11} for the entries in a contingency table; for example, n_{01} is the number of times that ice was forecast but no ice formed. Then the total cost over the period considered will have been $Ln_{10} + C(n_{01} + n_{11})$.

Figure 9 shows the economic cost associated with each of the seven forecasting methods as a function of the cost–loss ratio, R, when C equals 100,000 dollars. Since the costs for the bias-corrected forecast and the bias-corrected UWME are nearly identical to those for the raw forecast and the UWME, respectively, Figure 9 does not include results for the former two methods. Our probabilistic forecasts will have led to considerably lower costs than the raw forecast, at all thresholds. The naive forecast will have led to a total cost of 27.5 million dollars if $R < 0.07$, in which case it always predicts ice. At thresholds $R \geq 0.07$, the naive forecast never predicts ice and its total cost is $Ln_{10} = \frac{C}{R} n_{10}$.

To give a more detailed example, consider the case in which C equals 100,000 dollars and L equals 1,000,000 dollars, so that $R = 0.1$, which is a realistic assumption. In that case, acting on the basis of the raw forecast will have resulted in a total cost of 46.0 million dollars over the two winter seasons. Bias-correcting the raw forecast will have lowered the cost only slightly, to 45.9 million dollars. Acting on the basis of the UWME and bias-corrected UWME will have resulted in total costs of 31.1 and 30.5 million dollars, respectively. With the marginal and the spatial model, the cost will have been cut to

Table 2. Brier scores for probability forecasts of ice formation at individual sites and along the section $\mathcal{I}$ of I-90 for the 2003–2004 and 2004–2005 winter seasons. The uncertainty, reliability, and resolution components of the Brier score are shown in parentheses

	Ice at observation sites	Ice along $\mathcal{I}$
Raw forecast	0.179 (0.138; 0.070; 0.030)	0.280 (0.234; 0.085; 0.040)
Bias-corrected forecast	0.167 (0.138; 0.060; 0.031)	0.247 (0.234; 0.064; 0.051)
UWME	0.126 (0.138; 0.022; 0.035)	0.193 (0.234; 0.037; 0.078)
Bias-corrected UWME	0.152 (0.138; 0.047; 0.033)	0.203 (0.234; 0.045; 0.076)
Marginal model	0.103 (0.138; 0.043; 0.078)	0.167 (0.234; 0.103; 0.171)
Spatial model	0.115 (0.138; 0.045; 0.069)	0.163 (0.234; 0.080; 0.151)

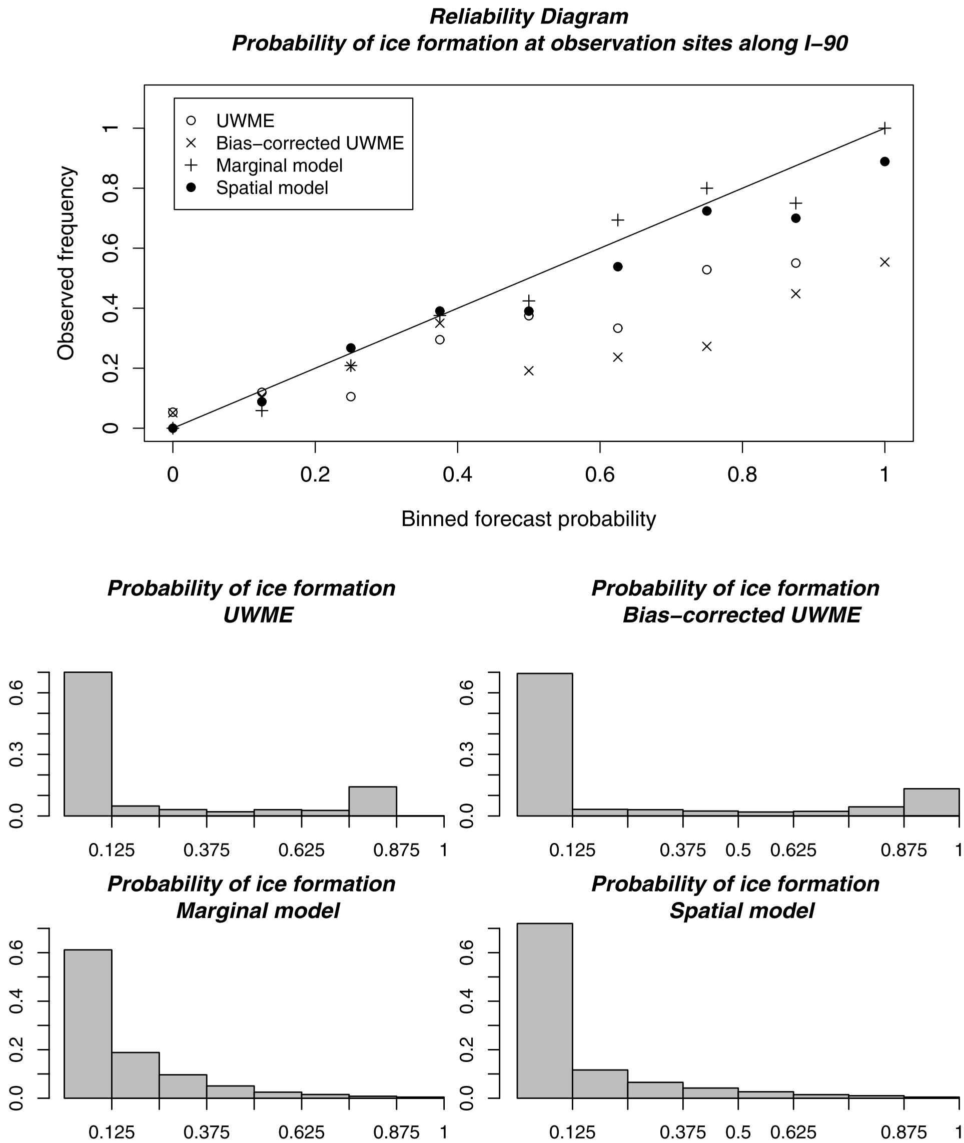


Figure 7. Reliability diagram for probability forecasts of ice formation at observation sites by the UWME, the bias-corrected UWME, the marginal model, and the spatial model, for the 2003–2004 and 2004–2005 winter seasons. Histograms of the forecast probabilities are also shown.

23.2 and 23.5 million dollars. Thus our results suggest that using our probabilistic forecasts can have lowered the overall economic loss by nearly 50%, or about 11 million dollars yearly on just one mountain pass, when compared to the raw forecast, and by 24% or 3.8 million dollars yearly, when compared to the ensemble forecasts. Bias removal alone reduces costs, but only slightly so.

Our test dataset consisted of 275 cases, so in our example with $C = 100{,}000$ dollars, taking anti-icing measures every day will have cost 27.5 million dollars. Basing decisions on the raw forecasts will have led to a loss considerably greater than this, showing the danger of relying on deterministic guidance when the costs and losses are as asymmetric as in this case. Using probabilistic forecasts of ice provided by the marginal model will have yielded a reduction in economic loss of about 15%, or 2 million dollars per year compared with the strategy of always anti-icing.

4. DISCUSSION

We develop two ways of estimating the probability of ice formation on a roadway. The simpler one ignores spatial dependence and the more complex one models spatial dependence explicitly. In our experiments, both probabilistic methods considerably outperformed the raw forecast and the ensemble forecasts we considered and will almost halve the total economic cost relative to relying on deterministic guidance. The two methods give similar results because much of the spatial dependence in ice formation was already accounted for by the numerical weather forecasts.

Our approach to the problem of predicting road ice goes beyond previous approaches by providing probabilistic forecasts rather than deterministic predictions. Additionally, it provides forecasts that are better calibrated and more reliable than probabilistic forecasts produced by ensembles of numerical forecasts.

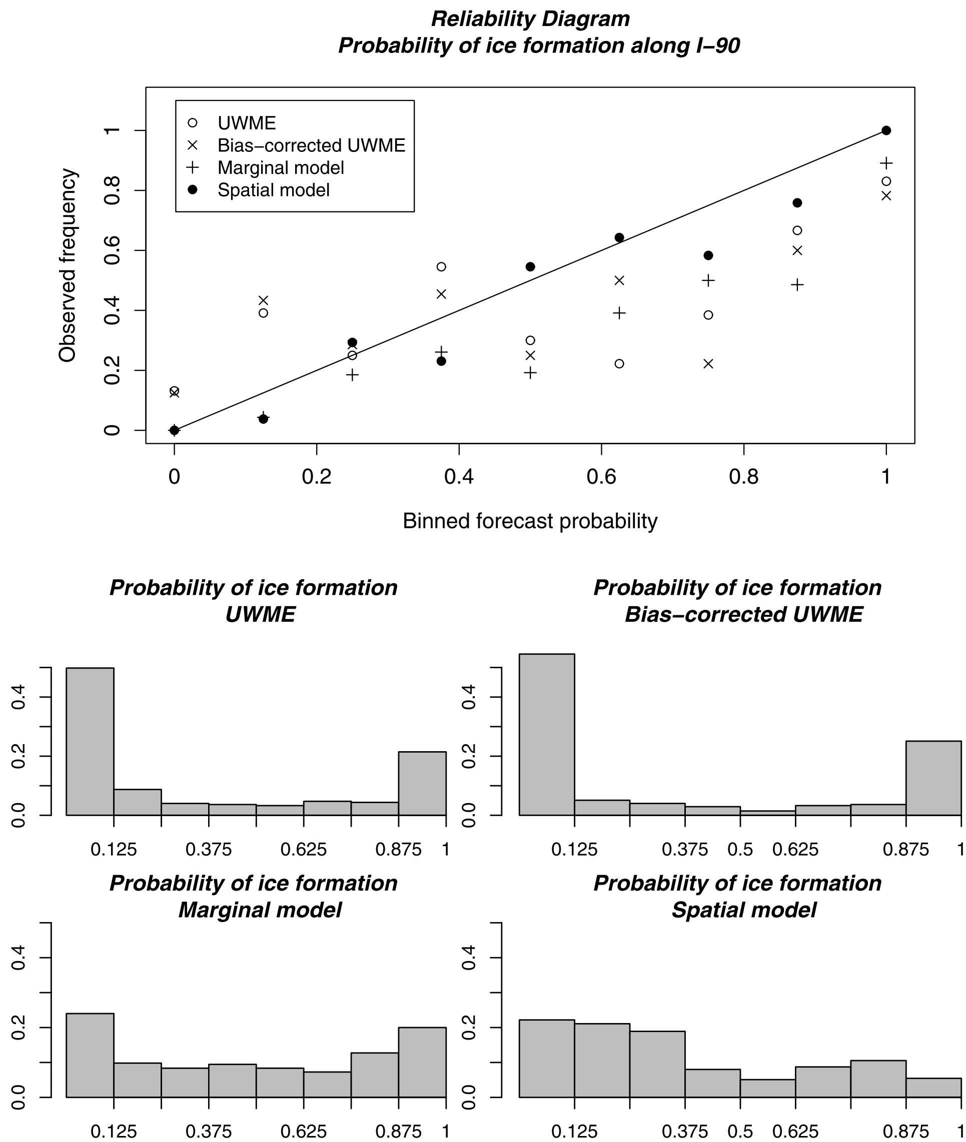


Figure 8. Reliability diagram for probability forecasts of ice formation along the section $\mathcal{I}$ of I-90 for the UWME, the bias-corrected UWME, the marginal model, and the spatial model for the 2003–2004 and 2004–2005 winter seasons. Histograms of the forecast probabilities are also shown.

Commonly, road ice is forecast using mathematical models that reproduce the physical interactions between the road and the atmosphere (Sass 1992; Shao and Lister 1995; Best 1998; Chapman, Thornes, and Bradley 2001b; Crevier and Delage 2001; Korotenko 2002). Such models take into account meteorological parameters, such as air temperature, precipitation, wind direction, wind speed, humidity, and dew point, and predict both road surface temperature and road conditions (Chapman, Thornes, and Bradley 2001b). However, despite the high level of detail, their predictions are not always accurate (Shao 1998). For example, three numerical road models used in the U.K. were found to be negatively biased (Chapman, Thornes, and Bradley 2001b).

Statistical methods are applied to select the weather and other variables that best predict road surface temperature (Shao and Lister 1996; Chapman, Thornes, and Bradley 2001a; Thornes, Cavan, and Chapman 2005). Similarly, statistical postprocessing methods for the outputs of numerical weather prediction or road prediction models are proposed. Shao (1998) used a backpropagation neural network to postprocess short-range forecasts of road surface temperature. Sherif and Hassan (2004) generated predictions of road surface temperature by linear regression of the observed pavement temperature on the numerical forecasts of selected meteorological variables. These methods are all deterministic in the sense that they are intended to yield point forecasts, rather than predictive distributions.

Our marginal and spatial models are fairly simple, yet they provide big improvements over the raw forecast and over ensembles of forecasts, both in terms of accuracy and economic value. The method for computing the predictive mean in our marginal model can be viewed as an instance of Model Output Statistics (MOS; Glahn and Lowry 1972; Wilks 2006). However, MOS was not previously applied to road ice forecasting and MOS does not yield probabilistic forecasts. Our spatial model can be viewed as a generalization of the model of Gel, Raftery, and Gneiting (2004) for probabilistic forecasting

Table 3. Forecasts and observations of ice formation for the naive forecast, the raw forecast, the bias-corrected forecast, the UWME, the bias-corrected UWME, the marginal model, and the spatial model at threshold probability $R = 0.1$ for the 2003–2004 and 2004–2005 winter seasons. Both the raw and the bias-corrected forecasts are deterministic, so their entries do not depend on the threshold probability

Forecast type	Ice forecast	Ice observed	Not observed
Naive forecast	Yes	0	0
	No	103	172
Raw forecast	Yes	68	42
	No	35	130
Bias-corrected forecast	Yes	67	32
	No	36	140
UWME	Yes	88	73
	No	15	99
Bias-corrected UWME	Yes	88	67
	No	15	105
Marginal model	Yes	101	111
	No	2	61
Spatial model	Yes	102	123
	No	1	49

of temperature fields to simultaneous forecasting of temperature and precipitation fields.

In our spatial model, precipitation occurrence is modeled by a latent Gaussian random field with spatial dependence between sites. Similar models for modeling precipitation were used by others (Bardossy and Plate 1992; Hughes and Guttorp 1994; Hutchinson 1995; Guillot 1999; Hughes, Guttorp, and Charles 1999; Sansò and Guenni 1999, 2000, 2004; Allcroft and Glasbey 2003; Rappold, Gelfand, and Holland 2008). Our main contribution here is modeling temperature and precipitation jointly to generate forecasts of ice.

One direction in which our approach can be expanded will be to make use of available forecast ensembles, perhaps using a Bayesian model averaging approach (Raftery et al. 2005; Berrocal, Raftery, and Gneiting 2007; Sloughter et al. 2007). Although our methods outperform ensemble forecasts, combining the two might work better than either one individually.

There are several other ways in which the two models presented in this article can be expanded. Some or all of the coefficients in both the spatial and marginal models for temperature and precipitation can be modeled as spatially varying parameters. In the spatial model, a more flexible class of covariance structures, such as the Matérn covariance function (Guttorp and Gneiting 2006), can be used to model the spatial dependence in the forecast errors of temperature and precipitation. However, it is not clear that this will lead to improvements in predictive performance since, as we noted, the numerical forecasts themselves already account for much of the spatial dependence in the data.

In estimating the model parameters, we have fit the two models for each day using a "sliding window" training period. This is a simple adaptive approach that, nevertheless, yields good results. A different approach will be to develop a dynamic version of the models presented here and estimate the parameters using all the data at once. We did not pursue this approach mainly for two reasons. First, our interest was in developing a method that can be operationally implemented by forecasters for real-time forecasting of road ice. Thus, we preferred a simple model that is easy to fit and interpret over a more complex one. Second, despite the spatio-temporal nature of the data, there was no need to model both the spatial and the temporal structure. Numerical weather forecasts are obtained by solving a system of partial differential equations that describe how the atmosphere evolves in time, and so they already incorporate temporal dependence. By including the numerical forecasts in the mean functions of the Gaussian processes for temperature and precipitation, we are accounting for temporal dependence. To confirm this, Figure 10 shows the empirical autocorrelation function of the raw forecast errors and the residuals from the regression model of Equation (2) at the Snoqualmie River Bridge station over three winter seasons. The results were similar for other stations.

To estimate the model parameters we adopt a two-stage approach, in which different parameters are estimated separately,

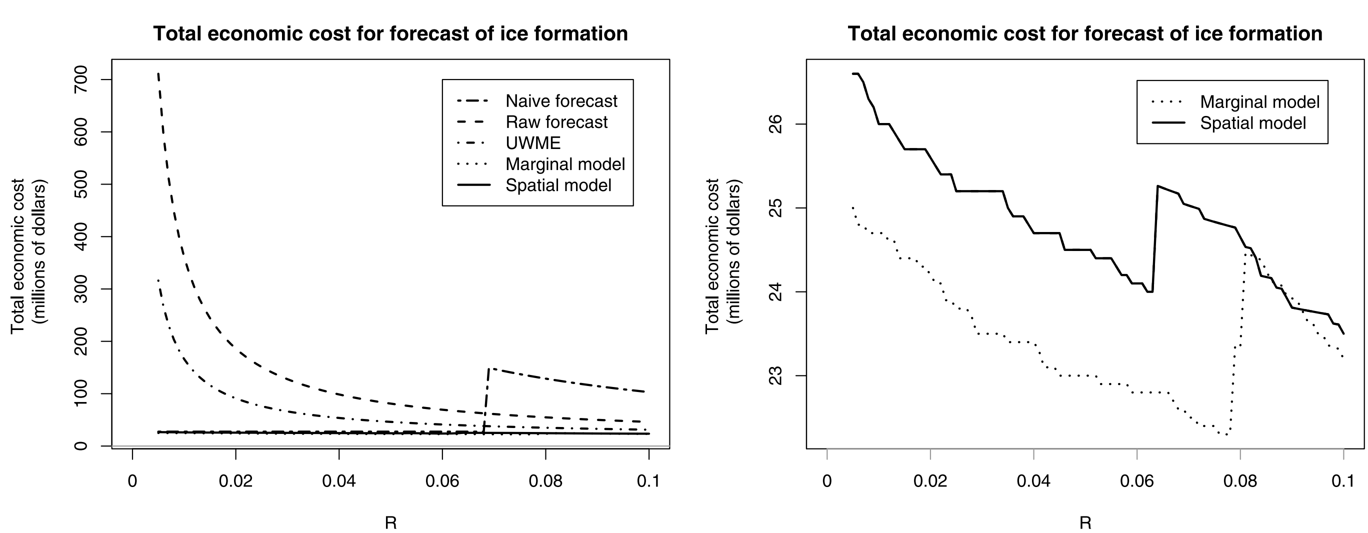


Figure 9. Total economic cost associated with forecasts of ice formation provided by the naive forecast, the raw forecast, the UWME, the marginal model, and the spatial model. The costs refer to the 2003–2004 and 2004–2005 winter seasons and depend on the cost–loss ratio, R.

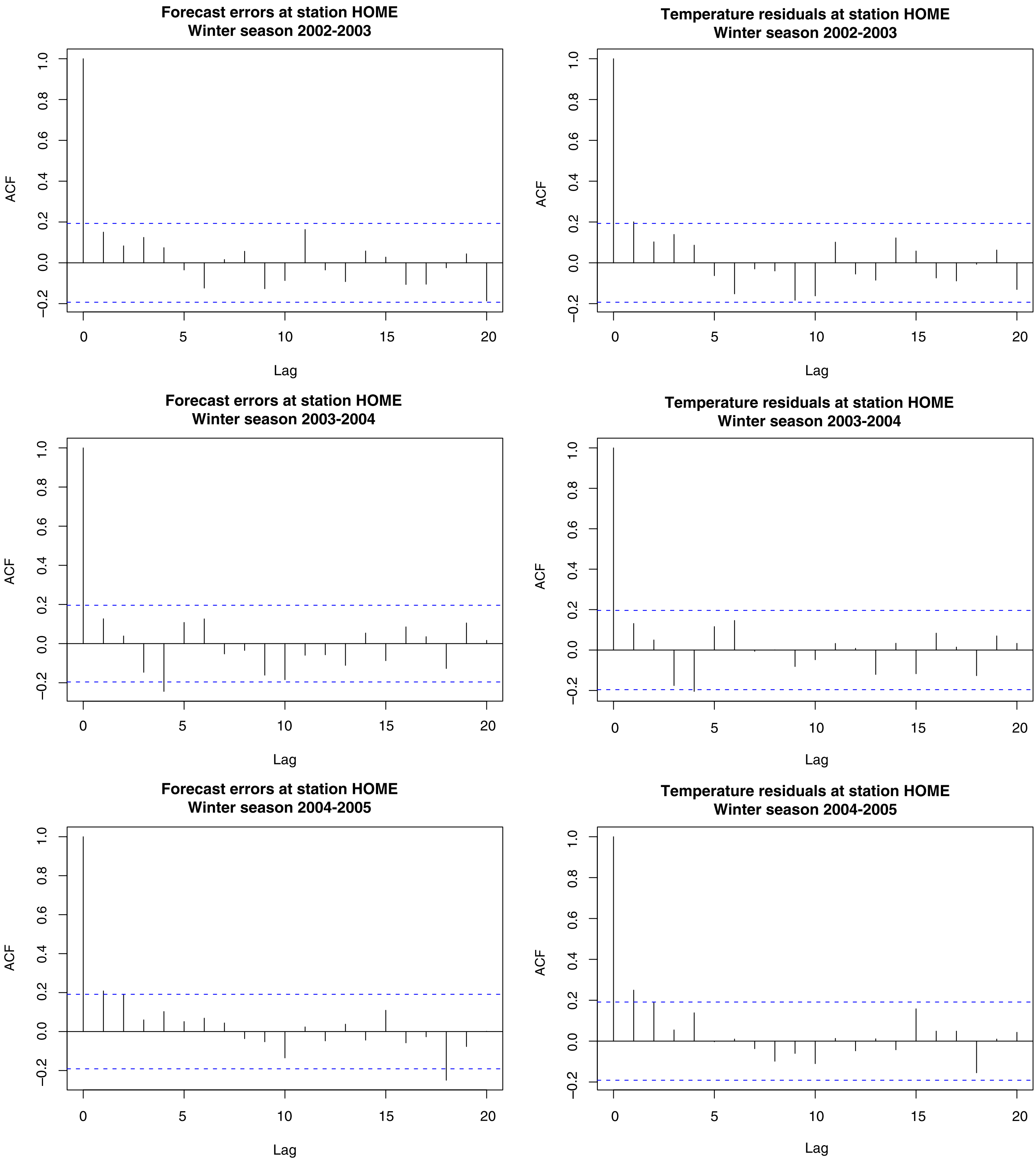


Figure 10. Empirical autocorrelation function for temperature forecast errors from the raw forecast and residuals from the regression model in Equation (2) at the Snoqualmie River Bridge (HOME) station for the winter seasons 2002–2003, 2003–2004, and 2004–2005. A color version of this figure is available in the electronic version of this article.

but that, as our results show, works well. Simultaneous estimation of all the parameters using Markov chain Monte Carlo often does not work well in this type of situation (Sahu, Gelfand, and Holland 2006), and might not be fast and simple enough for real-time implementation.

We developed our methods using numerical weather forecasts of air temperature and rainfall amount, and we defined ice as the joint occurrence of air temperature at or below freezing and precipitation. It will be more accurate to use observations of whether or not road ice actually occurred, but these were not available. Failing that, a more accurate definition of road ice will be the joint occurrence of pavement (rather than air) temperature at or below freezing and nonzero precipitation. Numerical forecasts of pavement temperature and observations of road surface temperature were not available to us. However, both the marginal model and the spatial model method can be adapted and applied to forecasts and observations of pavement temperature rather than air temperature.

[Received April 2007. Revised March 2009.]

REFERENCES

Albert, J. H., and Chib, S. (1993), "Bayesian Analyis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669–679. [526]

Allcroft, D. J., and Glasbey, C. A. (2003), "A Latent Gaussian Markov Random-Field Model for Spatiotemporal Rainfall Disaggregation," *Applied Statistics*, 52, 487–498. [526,534]

American Meteorological Society (2002), "Enhancing Weather Information With Probability Forecasts," *Bulletin of the American Meteorological Society*, 83, 450–452. [522]

Bardossy, A., and Plate, E. J. (1992), "Space–Time Model for Daily Rainfall Using Atmospheric Circulation Patterns," *Water Resources Research*, 28, 1247–1259. [526,534]

Berrocal, V. J., Raftery, A. E., and Gneiting, T. (2007), "Combining Spatial Statistical and Ensemble Information in Probabilistic Weather Forecasts," *Monthly Weather Review*, 135, 1386–1402. [534]

Best, M. J. (1998), "A Model to Predict Surface Temperatures," *Boundary-Layer Meteorology*, 88, 279–306. [533]

Brier, G. W. (1950), "Verification of Forecasts Expressed in Terms of Probability," *Monthly Weather Review*, 78, 1–3. [528]

Chapman, L., Thornes, J. E., and Bradley, A. V. (2001a), "Modelling of Road Surface Temperature From a Geographical Parameter Database. Part 1: Statistical," *Meteorological Applications*, 8, 409–419. [522,523,533]

——— (2001b), "Modelling of Road Surface Temperature From a Geographical Parameter Database. Part 2: Numerical," *Meteorological Applications*, 8, 421–436. [533]

Chilès, J.-P., and Delfiner, P. (1999), *Geostatistics: Modeling Spatial Uncertainty*, New York: Wiley. [526]

Cressie, N. A. C. (1993), *Statistics for Spatial Data* (Revised ed.), New York: Wiley. [526,528]

Crevier, L. P., and Delage, Y. (2001), "METRo: A New Model for Road-Condition Forecasting in Canada," *Journal of Applied Meteorology*, 40, 2026–2037. [533]

Eckel, F. A., and Mass, C. F. (2005), "Aspects of Effective Mesoscale, Short-Range Ensemble Forecasting," *Weather and Forecasting*, 20, 328–350. [523,530]

Eriksson, M., and Norrman, J. (2001), "Analysis of Station Locations in a Road Weather Information System," *Meteorological Applications*, 8, 437–448. [522]

Federal Highway Administration (2006), "Best Practices for Road Weather Management. Version 2.0," available at *http://ops.fhwa.dot.gov/Weather/best_practices/CaseStudies/029.pdf*. [522]

Gel, Y., Raftery, A. E., and Gneiting, T. (2004), "Calibrated Probabilistic Mesoscale Weather Field Forecasting: The Geostatistical Output Perturbation (GOP) Method" (with discussion), *Journal of the American Statistical Association*, 99, 575–588. [533]

Gelfand, A. E., and Smith, A. F. M. (1990), "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 398–409. [528]

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004), *Bayesian Data Analysis* (2nd ed.), Boca Raton, FL: Chapman & Hall/CRC. [528]

Glahn, H. R., and Lowry, D. A. (1972), "The Use of Model Output Statistics (MOS) in Objective Weather Forecasting," *Journal of Applied Meteorology*, 11, 1203–1211. [533]

Gneiting, T., and Raftery, A. E. (2005), "Weather Forecasting Using Ensemble Methods," *Science*, 310, 248–249. [522]

Grell, G. A., Dudhia, J., and Stauffer, D. R. (1994), "A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5)," Technical Note NCAR/TN-398+STR, NCAR, MMM Division. [523]

Grimit, E. P., and Mass, C. F. (2002), "Initial Results of a Mesoscale Short-Range Ensemble Forecasting System Over the Pacific Northwest," *Weather and Forecasting*, 17, 192–205. [523]

Guillot, G. (1999), "Approximation of Sahelian Rainfall Fields With Meta-Gaussian Random Functions," *Stochastic Environmental Research and Risk Assessment*, 13, 100–112. [534]

Guttorp, P., and Gneiting, T. (2006), "Studies in the History of Probability and Statistics XLIX: On the Matérn Correlation Family," *Biometrika*, 93, 989–995. [534]

Hagedorn, R., Hamill, T. M., and Whitaker, J. S. (2008), "Probabilistic Forecast Calibration Using ECMWF and GFS Ensemble Reforecasts. Part I: Two-Meter Temperatures," *Monthly Weather Review*, 136, 2608–2617. [529]

Hamill, T. M., Hagedorn, R., and Whitaker, J. S. (2008), "Probabilistic Forecast Calibration Using ECMWF and GFS ensemble Reforecasts. Part II: Precipitation," *Monthly Weather Review*, 136, 2620–2632. [529]

Handcock, M. S., and Wallis, J. R. (1994), "An Approach to Statistical Spatial–Temporal Modeling of Meteorological Fields," *Journal of the American Statistical Association*, 89, 368–378. [524]

Hughes, J. P., and Guttorp, P. (1994), "Incorporating Spatial Dependence and Atmospheric Data in a Model of Precipitation," *Journal of Applied Meteorology*, 33, 1503–1515. [534]

Hughes, J. P., Guttorp, P., and Charles, S. P. (1999), "A Non-Homogeneous Hidden Markov Model for Precipitation Occurrence," *Applied Statistics*, 48, 15–30. [534]

Hutchinson, M. F. (1995), "Stochastic Space–Time Weather Models From Ground-Based Data," *Agricultural and Forest Meteorology*, 73, 237–264. [526,534]

Jun, M., Knutti, R., and Nychka, D. (2008), "Spatial Analysis to Quantify Model Bias and Dependence: How Many Climate Models Are There?" *Journal of the American Statistical Association*, 103, 934–947. [523]

Korotenko, K. (2002), "An Automated System for Prediction of Icing on the Road," in *Proceedings of International Conference on Computational Science—ICSS 2002, Part III. Lecture Notes in Computer Science*, Vol. 2331, Berlin: Springer, pp. 1193–1200. [533]

Mass, C., Albright, M., Ovens, D., Steed, R., MacIver, M., Grimit, E., Eckel, T., Lamb, B., Vaughan, J., Westrick, K., Storck, P., Colman, B., Hill, C., Maykut, N., Gilroy, M., Ferguson, S. A., Yetter, J., Sierchio, J. M., Bowman, C., Stender, R., Wilson, R., and Brown, W. (2003), "Regional Environmental Prediction Over the Pacific Northwest," *Bulletin of the American Meteorological Society*, 84, 1353–1366. [523]

Murphy, A. H. (1973), "A New Vector Partition of the Probability Score," *Journal of Applied Meteorology*, 12, 595–600. [530]

National Research Council of the National Academies (2006), *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts*, Washington: The National Academies Press. [522]

Norrman, J., Eriksson, M., and Lindqvist, S. (2000), "Relationships Between Road Slipperiness, Traffic Accident Risk and Winter Road Maintenance Activity," *Climate Research*, 15, 185–193. [522]

Palmer, T. N. (2000), "Predicting Uncertainty in Forecasts of Weather and Climate," *Reports on Progress in Physics*, 63, 71–116. [522]

Paulson, T. (2001), "In Avalanche Country, Thinnest of Defenses Hangs Tough," *Seattle Post-Intelligencer*, December 21, 2001. Available at *http://www.seattlepi.com/local/52258_avalanche27.shtml*. [531]

Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005), "Using Bayesian Model Averaging to Calibrate Forecast Ensembles," *Monthly Weather Review*, 133, 1155–1174. [529,534]

Ramakrishna, D. M., and Viraraghavan, T. (2005), "Environmental Impact of Chemical Deicers—A Review," *Water, Air, and Soil Pollution*, 166, 49–63. [522]

Rappold, A., Gelfand, A. E., and Holland, D. M. (2008), "Modeling Mercury Deposition Through Latent Space–Time Processes," *Journal of the Royal Statistical Society, Ser. C*, 57, 187–205. [526,534]

Richardson, D. S. (2000), "Skill and Relative Economic Value of the ECMWF Ensemble Prediction System," *Quarterly Journal of the Royal Meteorological Society*, 126, 649–667. [522]

Roulston, M. S., Bolton, G. E., Kleit, A. N., and Sears-Collins, A. L. (2006), "A Laboratory Study of the Benefits of Including Uncertainty Information in Weather Forecasts," *Weather and Forecasting*, 21, 116–122. [522]

Sahu, S. K., Gelfand, A. E., and Holland, D. M. (2006), "Spatio-Temporal Modeling of Fine Particulate Matter," *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 61–86. [526,535]

Sansò, B., and Guenni, L. (1999), "Venezuelan Rainfall Data Analysed by Using a Bayesian Space–Time Model," *Applied Statistics*, 48, 345–362. [526,534]

——— (2000), "A Nonstationary Multisite Model for Rainfall," *Journal of the American Statistical Association*, 95, 1089–1100. [526,534]

——— (2004), "A Bayesian Approach to Compare Observed Rainfall Data to Deterministic Simulations," *Environmetrics*, 15, 597–612. [526,534]

Sass, B. H. (1992), "A Numerical Model for Prediction of Road Temperature and Ice," *Journal of Applied Meteorology*, 31, 1499–1506. [522,533]

Shao, J. (1998), "Improving Nowcasts of Road Surface Temperature by a Backpropagation Neural Network," *Weather and Forecasting*, 13, 164–171. [522,533]

Shao, J., and Lister, P. J. (1995), "The Prediction of Road Surface-State and Simulations of the Shading Effect," *Boundary-Layer Meteorology*, 73, 411–419. [533]

——— (1996), "An Automated Nowcasting Model of Road Surface Temperature and State for Winter Road Maintenance," *Journal of Applied Meteorology*, 35, 1352–1361. [522,523,533]

Shao, X., Stein, M., and Ching, J. (2007), "Statistical Comparisons of Methods for Interpolating the Output of a Numerical Air Quality Model," *Journal of Statistical Planning and Inference*, 137, 2277–2293. [523]

Sherif, A., and Hassan, Y. (2004), "Modelling Pavement Temperature for Winter Maintenance Operations," *Canadian Journal of Civil Engineering*, 31, 369–378. [522,533]

Sloughter, J. M., Raftery, A. E., Gneiting, T., and Fraley, C. (2007), "Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging," *Monthly Weather Review*, 135, 3209–3220. [524,529,534]

Thornes, J. E., Cavan, G., and Chapman, L. (2005), "XRWIS: The Use of Geomatics to Predict Winter Road Surface Temperatures in Poland," *Meteorological Applications*, 12, 83–90. [523,533]

Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables," *Econometrica*, 26, 24–36. [526]

Wilks, D. S. (2006), *Statistical Methods in the Atmospheric Sciences* (2nd ed.), San Diego: Academic Press. [533]