Diurnal Corrections of Short-Term Surface Temperature Forecasts Using the Kalman Filter

MARIKEN HOMLEID

Norwegian Meteorological Institute, Oslo, Norway

(Manuscript received 22 July 1994, in final form 2 May 1995)

ABSTRACT

Surface temperature forecasts from numerical weather prediction models are known to have systematic errors, partly due to poor resolution of the topography and deficiencies in the physical formulation. The Kalman filter theory provides an excellent tool for combining observations and forecasts to correct such systematic errors. As detailed in this paper, deviations between temperature forecasts and observations were studied at several locations in Norway. It was found that the sign and magnitude of the deviations varied considerably in both time and space. Moreover, the diurnal variation of the systematic errors increased as the diurnal variation between daytime and nighttime increased. A Kalman filter model has been constructed that allows for diurnally varying corrections. Eight corrections, valid at 0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC, are calculated simultaneously. These corrections, which are weighted means of previous differences between observations and forecasts, are applied to correct +3-, +6-, +9-, . . . , +48-h temperature forecasts. One set of corrections is calculated and updated for each of about 240 observing stations throughout Norway. The results of the correction procedure have been verified for about one-half of them. It was found that the correction procedure reduced the monthly bias of the forecasts at each observing station to values close to zero. The monthly standard deviations of the differences between forecasts and observations remained essentially unchanged. They were slightly reduced in summertime but could be slightly increased in wintertime for the long-term forecasts.

1. Introduction

Surface temperature forecasts from the numerical weather prediction (NWP) limited area model, LAM50, at the Norwegian Meteorological Institute (DNMI) are presented for some users on "meteograms." The LAM50 model (Bratseth 1986; Grønås and Midtbø 1987) is run every six hours. The meteograms are also produced every six hours for about 900 locations throughout Norway and some major European cities, giving surface forecasts of temperature, pressure, wind, precipitation, and clouds every three hours from 0 to 42 or 48 hours ahead. An example of a meteogram produced at 0000 UTC 4 July 1994 is shown in Fig. 1. This meteogram is presented with two temperature curves, one that is direct model output and one that is corrected by the Kalman filter procedure that will be presented in this paper. The surface pressure, the wind, and the cloud amount forecasts are all direct model output.

This paper will focus on the forecasts of temperatures 2 m above surface, T_{2m}, at about 240 meteogram locations that have corresponding surface observations. When comparing temperature forecasts and observations, both systematic and unsystematic differences are found. The systematic errors are partly due to the resolution of the NWP model that cannot resolve subgrid phenomena and that introduces differences between the topography as seen by the NWP model and the real topography. The difference between the topographical height in the LAM50 (Fig. 2) and the actual terrain height might be several hundred meters, implying large and systematic errors in the T_{2m} forecasts. Deficient physical formulations and the time resolution of the NWP model probably also contribute to the systematic errors in the temperature forecasts. The Kalman filter theory (Kalman 1960, 1963; Kalman and Bucy 1961; Gelb 1974; Priestley 1981) provides an excellent tool for combining observations and forecasts to correct for such systematic errors. General linear Kalman filter theory is presented briefly in section 2, concentrating on the Kalman filter equations. As compared to other methods for correcting NWP model output, such as the traditional model output statistics (MOS) approach (Glahn and Lowry 1972), the Kalman filter approach benefits from updating the corrections or the regression coefficients recursively, allowing it to adapt when there is a major change in the NWP model. The updatable MOS technique (Ross 1989) updates the regression coefficients each month, also reducing the problem of frequent changes in the NWP model. A large database containing forecasts and

Corresponding author address: Dr. Mariken Homleid, DNMI, Norwegian Meteorological Institute, P.O. Box 43 Blindern, 0313 Oslo, Norway.

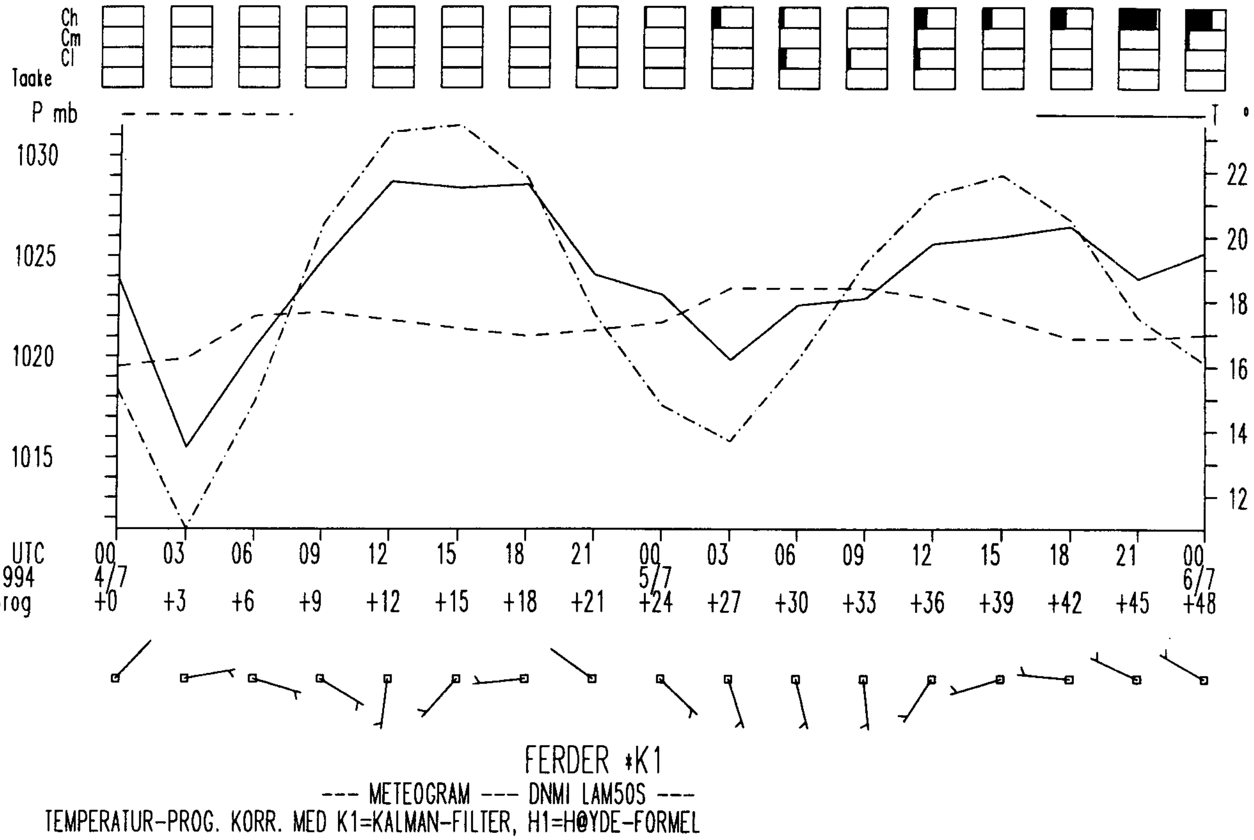

FIG. 1. "Meteogram" for Færder with forecasts from LAM50 of surface temperature, both uncorrected (dash-dot line), and Kalman filter corrected (full line); the scale is to the right. The pressure forecasts are given by the dashed line, with the scale to the left. The cloud amounts are given in the boxes above, with Ch, Cm, Cl, and Taake representing high, medium, and low clouds, and fog. The wind forecasts are given by the arrows below. The forecasts are produced at 0000 UTC 4 July 1994 and are valid 2 days ahead.

observations is very useful when defining the Kalman filter model, in particular when taking into account other weather parameters, but it is not a prerequisite for applying the Kalman filter correction procedure. Several papers describe Kalman filter models for adaptive correction of surface temperature forecasts (Persson 1991; Simonsen 1991; Kilpinen 1992). They take account of the flexibility of the Kalman filter by allowing for other parameters such as relative humidity and wind to explain the differences between forecasted and observed temperatures. Since Norway is a country with large variations in weather and topography, the inclusion of other weather parameters in a proper way would require careful studies of weather dependencies of the forecast errors at several stations representing different climatic and topographical conditions.

The Kalman filter model presented in this paper will only use temperature information to correct the forecasts. The purpose has been, as a first step, to construct a Kalman filter model that is general enough to be applied at all the 240 observing stations. The Kalman filter model is defined to simultaneously estimate eight corrections valid at 0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC. Section 3 illustrates the need for a temperature correction procedure by presenting bias characteristics of the LAM50 at several locations. In section 4, the Kalman filter model for estimating diurnal corrections of temperature forecasts is described. This model has been running operationally at DNMI since September 1992 at about 240 stations. The operational implementation is described briefly in section 5. How the Kalman filter works is explained in section 6 by some examples and in section 7 by summary statistics. A summary and discussion is given in section 8.

2. General linear Kalman filter theory

The Kalman filter theory provides equations for recursively updating estimates of an unknown process, combining observations related to the process and knowledge about the time evolution of the process. These equations were first derived within the context of "linear systems," and their applications were found within the field of navigation of satellites. Later, the Kalman filter formulation found an extremely wide

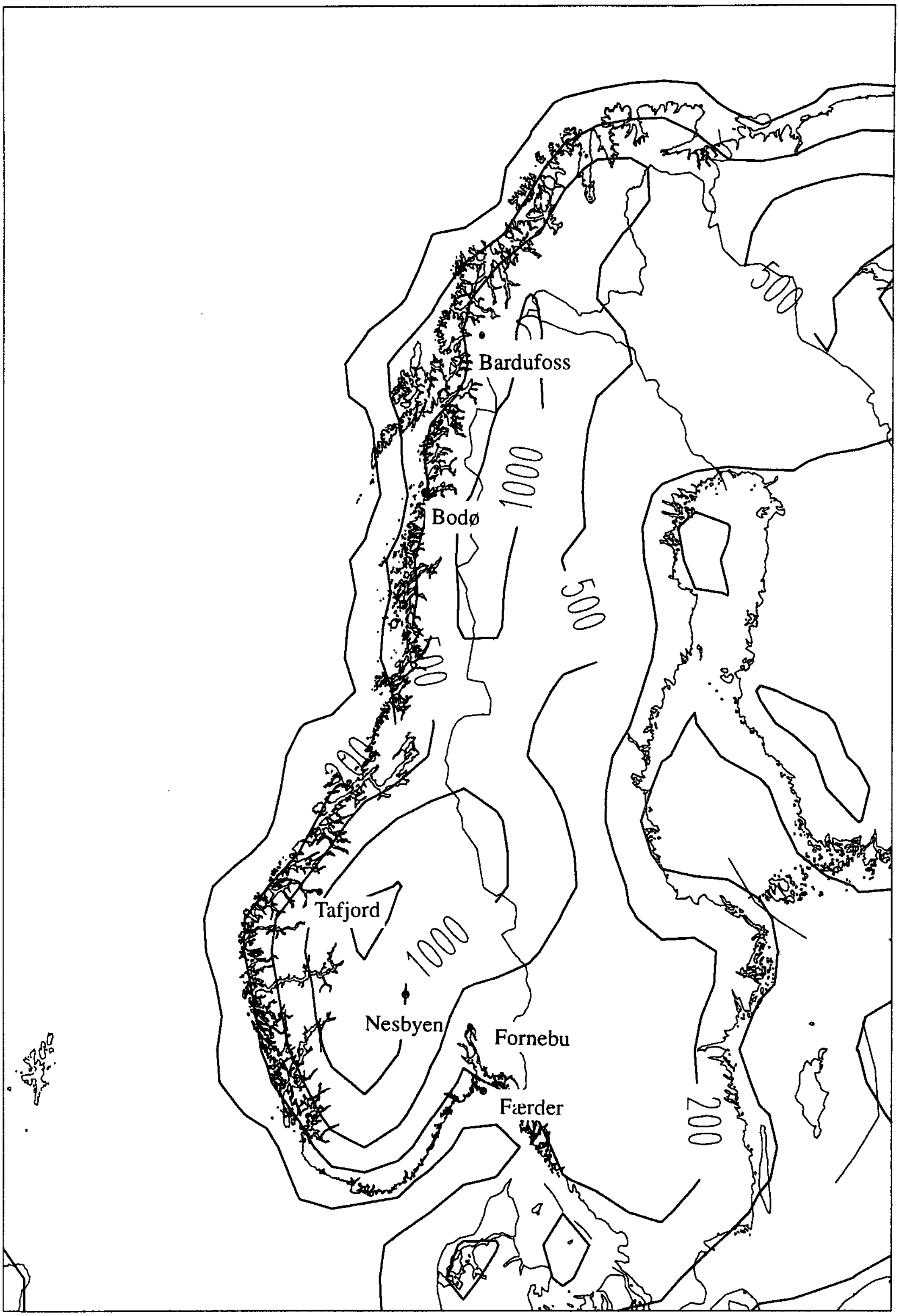

FIG. 2. Topographical height of LAM50 over Norway. The locations focused on in this study are displayed.

TABLE 1. Geographical position, terrain height, and topographical height of LAM50 (in meters) for the six locations focused on in this study.

	Geo. position	$H_{terrain}$	H_{LAM50}
Fornebu	59.54°N, 10.38°E	17	332
Færder	59.02°N, 10.32°E	8	154
Nesbyen	60.34°N, 9.08°E	167	1009
Tafjord	62.14°N, 7.25°E	30	870
Bodø	67.16°N, 14.22°E	13	443
Bardufoss	69.03°N, 18.33°E	79	700

range of statistical applications, for example, in Bayesian forecasting (Harrison and Stevens 1976), in the analysis of time series models with time-dependent coefficients (Young 1974; Bohlin 1976), and in general regression models (O'Hagan 1978).

The Kalman filter theory is too extensive to be thoroughly presented here. A complete description of the Kalman filter model and the derivation of equations giving optimal estimates of the unknown process can be found in Gelb (1974) or Priestley (1981). Given below are the basic equations describing the time evolution of the process and the relationship of the observations to the process and the recursive Kalman filter equations giving estimates of the unknown process. The main model assumptions are also stated.

Let $\mathbf{x}_t$ be a vector describing the state of the unknown process at time t, in our case the systematic deviation between observed and forecasted temperature. The state at time t is related to the state at time $(t-1)$ through the

system equation:

$$\mathbf{x}_t = \mathbf{F}_t\mathbf{x}_{t-1} + \mathbf{w}_t \quad t = 1, \cdots, T, \tag{2.1}$$

where $\mathbf{F}_t$ is the system matrix, and $\mathbf{w}_t$ is a vector giving the random change from time $(t-1)$ to time t.

The state $\mathbf{x}_t$ is not observable but is related to observations $\mathbf{y}_t$ (which in our case will be differences between forecasts and observations) through the

observation equation:

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{v}_t \quad t = 1, \cdots, T, \tag{2.2}$$

where $\mathbf{H}_t$ is the observation matrix, and $\mathbf{v}_t$ is a vector containing the random observation error.

The system matrix $\mathbf{F}_t$ and observation matrix $\mathbf{H}_t$ can vary in time but are supposed to be known. Terms $\mathbf{w}_t$ and $\mathbf{v}_t$ are Gaussian zero mean white noise processes, which means that $\mathbf{w}_t$ and $\mathbf{v}_t$ for all t have Gaussian distributions with zero means and that the noise is time independent; $E(\mathbf{w}_s \cdot \mathbf{w}_t^T) = 0$ and $E(\mathbf{v}_s \cdot \mathbf{v}_t^T) = 0$ for all s and t, $s \neq t$. Moreover, $\mathbf{w}_s$ and $\mathbf{v}_t$ are independent; $E(\mathbf{w}_s \cdot \mathbf{v}_t^T) = 0$ for all s and t. $\mathbf{W}_t = E(\mathbf{w}_t \cdot \mathbf{w}_t^T)$ is the covariance matrix of $\mathbf{w}_t$ and $\mathbf{V}_t = E(\mathbf{v}_t \cdot \mathbf{v}_t^T)$ is the covariance matrix of $\mathbf{v}_t$.

The Kalman filter theory gives an algorithm for recursively estimating the unknown state $\mathbf{x}_t$ utilizing all observations up to time t; $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_t$. The quality of the state estimates can be evaluated through estimates of their error covariance matrix $\mathbf{P}_t$, which are also given. The updating of the estimates from time $(t-1)$ to t is performed in two steps, and two types of estimates are defined:

- $\hat{\mathbf{x}}_{t|t-1}$ estimate of $\mathbf{x}_t$ given observations up to time $(t-1)$,
- $\hat{\mathbf{P}}_{t|t-1}$ estimate of the covariance matrix of the error of $\hat{\mathbf{x}}_{t|t-1}$,
- $\hat{\mathbf{x}}_{t|t}$ estimate of $\mathbf{x}_t$ given observations up to time t,
- $\hat{\mathbf{P}}_{t|t}$ estimate of the covariance matrix of the error of $\hat{\mathbf{x}}_{t|t}$.

Suppose the process is observed up to time $(t-1)$ and that the optimal estimate $\hat{\mathbf{x}}_{t-1|t-1}$ is computed. Using only observations up to time $(t-1)$, the best estimate of $\mathbf{x}_t$ is

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t\hat{\mathbf{x}}_{t-1|t-1}. \tag{2.3}$$

The covariance matrix of the observation error in $\hat{\mathbf{x}}_{t|t-1}$ is estimated by

$$\hat{\mathbf{P}}_{t|t-1} = \mathbf{F}_t\hat{\mathbf{P}}_{t-1|t-1}\mathbf{F}_t^T + \mathbf{W}_t. \tag{2.4}$$

When the observation at time t $\mathbf{y}_t$ becomes available, the estimate of $\mathbf{x}_t$ can be updated by forming a linear combination of the previous estimate $\hat{\mathbf{x}}_{t|t-1}$ and the prediction error $(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1})$:

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}), \tag{2.5}$$

where the "Kalman gain" matrix is

$$\mathbf{K}_t = \hat{\mathbf{P}}_{t|t-1}\mathbf{H}_t^T(\mathbf{H}_t\hat{\mathbf{P}}_{t|t-1}\mathbf{H}_t^T + \mathbf{V}_t)^{-1}. \tag{2.6}$$

The covariance of the error in $\hat{\mathbf{x}}_{t|t}$ is estimated by

$$\hat{\mathbf{P}}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\hat{\mathbf{P}}_{t|t-1}. \tag{2.7}$$

This set of equations is sufficient for updating the Kalman filter from time $(t-1)$ to t.

Before running the Kalman filter, an appropriate model has to be specified. Model specification involves setting up the system matrix $\mathbf{F}_t$ and the observation matrix $\mathbf{H}_t$ and estimating the covariance matrices of the system and observation noise $\mathbf{W}_t$ and $\mathbf{V}_t$. Estimates of the initial state $\hat{\mathbf{x}}_{0|0}$ and $\hat{\mathbf{P}}_{0|0}$ are also required before running the filter.

3. Systematic errors in short-term surface temperature forecasts

In order to emphasize the need for a correction scheme that adjusts the explicit temperature forecasts from the NWP model, and in order to lay the framework for the specification of an appropriate Kalman filter model, several figures are presented that illustrate how the temperature biases vary with respect to location, time of day, and month of year. The geographical positions of the stations examined are given in Table 1 and displayed in Fig. 2.

As an example, short-term temperature forecasts and observations at Bardufoss for March 1993 are presented in Fig. 3. The short-term forecasts valid at 0000, 0600, 1200, and 1800 UTC are +6-h forecasts issued six hours earlier. The short-term forecasts valid at 0300, 0900, 1500, and 2100 UTC are +3-h forecasts issued three hours earlier. From 1 to 19 March the forecasts are mostly too cold, except for short periods with very low temperature observations. In the stable period from 20 to 30 March, there is a systematic diurnal variation in the differences. The NWP model seems to be unable to describe the low temperatures typical in wintertime during periods with clear sky.

Monthly mean values of temperature observations and forecasts for each month in 1993 are compared in Fig. 4 for each of six stations in this study. The observation averages are based on all available observations, taken every three hours at Færder, Fornebu, Bodø, and Bardufoss, at 0000, 0600, 1200, and 1800 UTC at Tafjord and at 0600, 1200, 1500, and 1800 UTC at Nesbyen. The forecast mean values are averages of +3, +6, +9, +12, +15, +18, +21, and +24 h generated at 0000, 0600, 1200, and 1800 UTC, one line for each forecast issue time. Climatological mean values based on observations from 1961 to 1990 are also given. The figures reveal that the differences between monthly mean values of temperature observations and forecasts fluctuate dramatically from one station to another and throughout the year and that the values based on forecasts issued at 0000, 0600, 1200, and 1800 UTC are much the same. The large positive biases in the forecasts for Nesbyen and Bardufoss in December 1993 point to the difficulty of LAM50 describing very low temperatures typical at locations such as Nesbyen and Bardufoss in wintertime.

The variation in monthly mean values has also been studied as a function of forecast valid time. Results for Færder and Bardufoss are presented in Figs. 5 and 6, respectively. The figures clearly illustrate the diurnal variation in the systematic deviations between forecasts and observations and suggest that the expected forecast errors do not vary systematically as functions of forecast issue time and projection. The diurnal variation changes throughout the year, with largest amplitudes when the largest differences between day and night occur. In the southern part of Norway (e.g., Færder) the differences are largest from March to September. In the northern part (e.g., Bardufoss), with midnight sun during summer and no sun during winter, the differences are largest in March, April, September, and October. Results for Færder indicate that the temperature forecasts are too high during the day and too cold during the night. The observed temperatures are close to the sea temperature and do not follow the diurnal cycle of the NWP model. As remarked before, the opposite is the case for Bardufoss. The diurnal cycle in the temperature forecasts is smaller than what is observed.

The main features of the systematic deviations between temperature forecasts and observations deduced from monthly mean values can be summarized as follows: the systematic deviations vary with valid time of day throughout the year and from station to station. In periods with a strong diurnal cycle in the observed

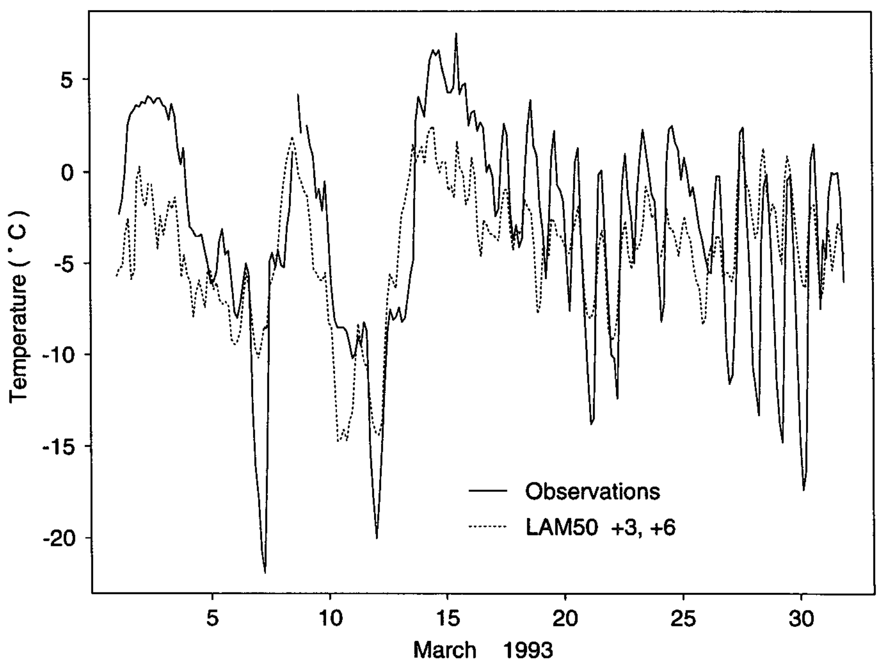

FIG. 3. Temperature observations (full line) and short-term (+3, +6 h) forecasts (dotted line), plotted every three hours at Bardufoss for March 1993.

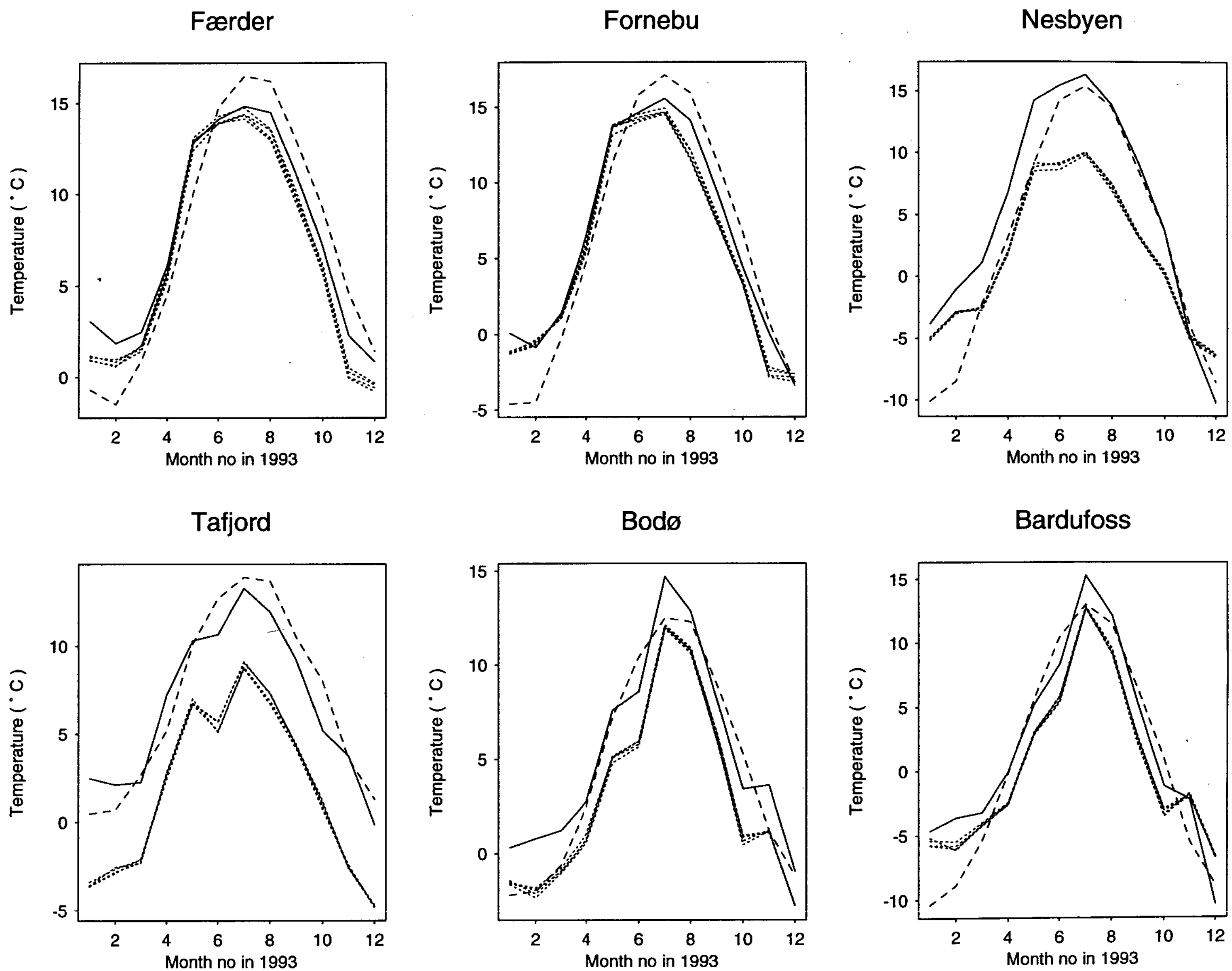

FIG. 4. Monthly mean values of temperature observations (full lines), forecasts (+3, +6, . . . , +24 h) issued at 0000, 0600, 1200, and 1800 UTC (dotted lines), and climatological mean values (1961–1990) (dashed lines) from six locations in Norway.

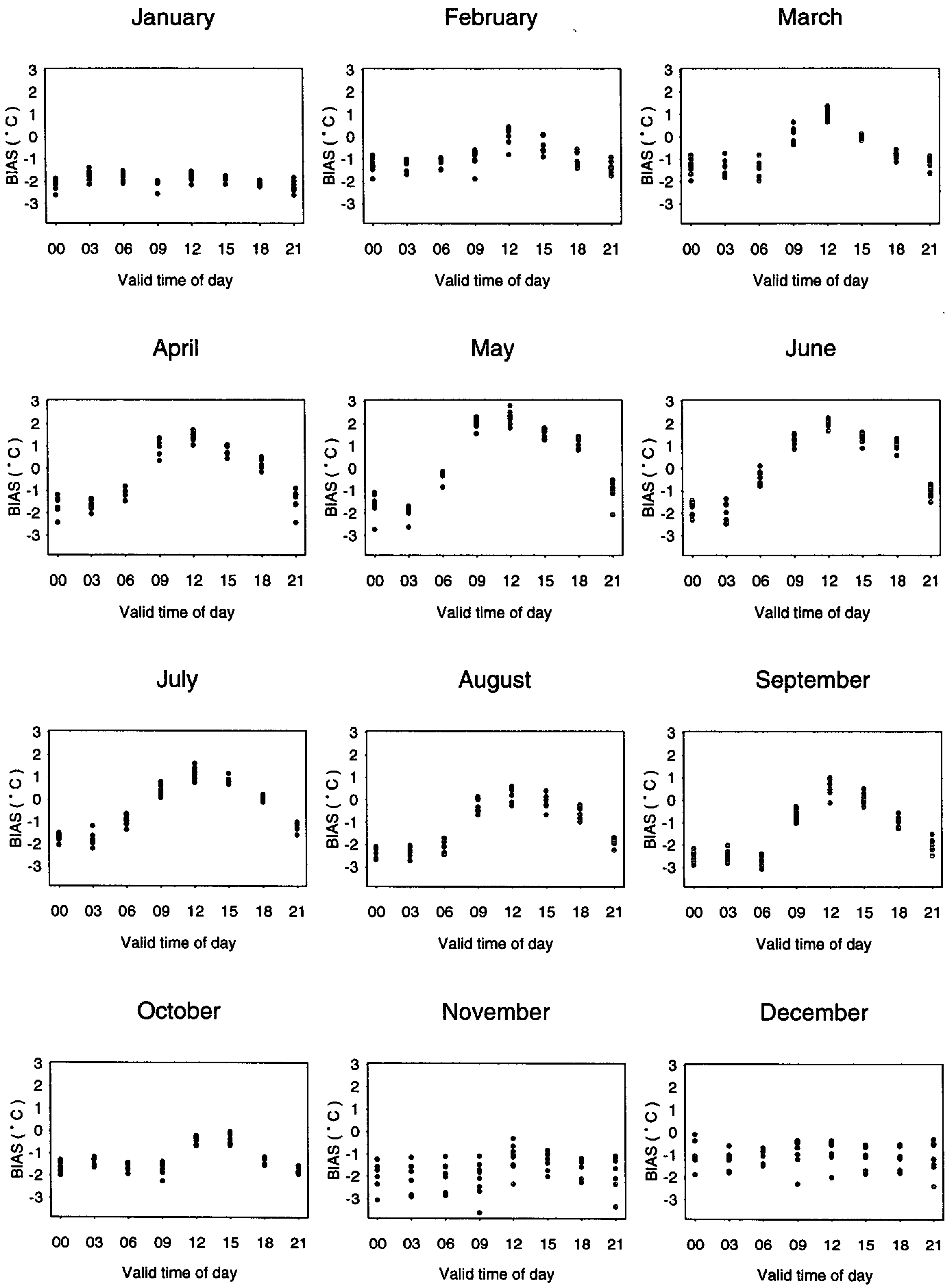

FIG. 5. Monthly mean values of differences between temperature forecasts and observations (BIAS) as a function of valid time of day. Each dot represents one specific forecast issue time and projection. The results are for Færder 1993.

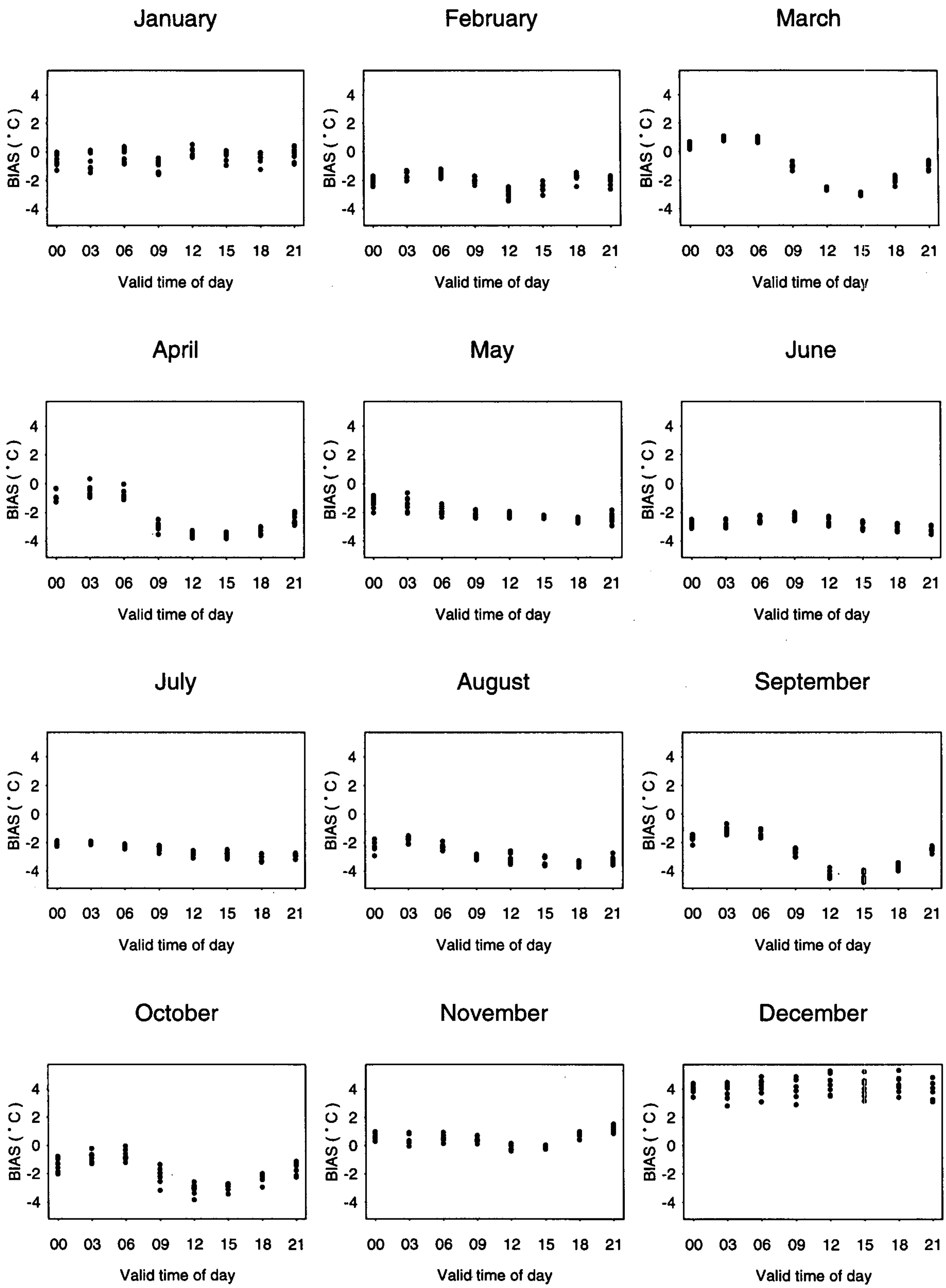

FIG. 6. As in Fig. 5 but for Bardufoss 1993.

temperature, there are frequently corresponding variations in the differences between the forecasts and the observations. The differences seem to be independent of forecast issue time and projection.

4. A Kalman filter model for correcting surface temperature forecasts

Norway is a country with large variations in weather, as illustrated by the studies of temperature forecasts and observations for the six sites above. The aim is to design a Kalman filter model for correcting temperature forecasts that is applicable at all 240 observation stations throughout the year. There are common features in the systematic differences between forecasts and observations that underlie the model specifications, in that for shorter or longer periods there are systematic diurnal variations in the differences. Defining the Kalman filter model involves setting up the system matrix $\mathbf{F}_t$ and the observation matrix $\mathbf{H}_t$ and specifying the statistical parameters $\mathbf{W}_t$ and V_t. Given below are the Kalman filter equations, a description of the Kalman filter procedure, and a discussion of the parameter specification.

System equation:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t. \tag{4.1}$$

The vector $\mathbf{x}_t$ includes eight elements $x_{i,t}$ that define the systematic deviations between the temperature forecasts and the observations at the observing times 0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC. The time step of the filter is three hours. The changes of the systematic deviations from time $(t-1)$ to t are assumed to be random, implying that the system matrix $\mathbf{F}_t$ is taken to be the identity matrix. The vector $\mathbf{w}_t$ includes eight elements, $w_{i,t}$ being the random change in $x_{i,t}$ from time $(t-1)$ to t. Term $\mathbf{w}_t$ is assumed to be white noise with time-independent covariance matrix $\mathbf{W}_t = \mathbf{W}$.

Observation equation:

$$y_t = \mathbf{H}_t\mathbf{x}_t + v_t. \tag{4.2}$$

Here, y_t is the difference between the observation and a short-term forecast valid at time t. The forecasts applied at 0000, 0600, 1200, and 1800 UTC are +6-h forecasts issued six hours earlier, and the forecasts applied at 0300, 0900, 1500, and 2100 UTC are +3-h forecasts issued three hours earlier. The observation matrix $\mathbf{H}_t$, which in our case is a 1×8 vector, is constructed in such a way that y_t is a function only of the $x_{i,t}$-element for time t. The observation equation can also be written $y_t = x_{i,t} + v_t$. Term $\mathbf{H}_t$ varies with time t, for example, $\mathbf{H}_t = (0\,0\,0\,0\,1\,0\,0\,0)$ at $t \sim 1200$ UTC. Term v_t, the observation error, is assumed to be white noise with time-independent variance $V_t^2 = V^2$.

The present model yields estimates of the systematic deviations between the observations and the short-term forecasts as a weighted mean of previous differences between observations and forecasts, with decreasing weight backward in time. The estimated deviations $\hat{\mathbf{x}}_{t|t}$ will be applied as corrections to the forecasts that are produced at time t. Forecasts valid at the same time will be equally corrected. For example, the 0 + 3-h and 0 + 27-h forecasts will be corrected by $\hat{x}_{2,t|t}$. This model allows for diurnal varying corrections when there is a diurnal variation in the differences between forecasts and observations.

a. Specification of statistical parameters

Before running the Kalman filter, the covariance matrix of the system noise $\mathbf{W}$ and the variance of the observation error V^2 need to be specified. These parameters can be estimated by a statistical estimation procedure [e.g., the EM (Expectation Maximization) algorithm (Dempster et al. 1977)] or by "tuning" them to make the Kalman filter work as wanted. To evaluate different candidates for $\mathbf{W}$ and V^2, the results of the filtering have been studied. The EM algorithm gives estimates that maximize the likelihood of the observations. In our case, these estimates resulted in a Kalman filter adapting too slowly to new weather conditions. The reason might be the known weakness of the specified Kalman filter model, where the differences between forecasts and observations are supposed to change slowly. The parameters chosen for the operational model allow faster adaptation to new conditions. With the same set of parameters, good results are obtained for all observation stations, both the ones that observe eight times a day and the ones that have less than eight observations a day.

The $\mathbf{W}$ matrix is designed as follows: $\mathbf{W}$ is the 8×8 covariance matrix of the system error. Term $\mathbf{W}$ is symmetric implying that it contains 36 elements to be specified. This number is reduced by assuming that the corrections are correlated in a systematic way:

1) the variances, $W_{ii,t}^2$, of the random change in the deviation from time $(t-1)$ to t do not vary throughout the day or from day to day: $W_{ii,t}^2 = W^2$ for $i = 1, \cdots, 8$, and all t;

2) the correlation between two deviations is only dependent on the delay in time between them.

The covariance matrix $\mathbf{W}$ can be written as a product of the variance W^2 and a correlation matrix:

$$\mathbf{W} = W^2 \times \begin{bmatrix} 1 & r_3 & r_6 & r_9 & r_{12} & r_9 & r_6 & r_3 \\ r_3 & 1 & r_3 & r_6 & r_9 & r_{12} & r_9 & r_6 \\ r_6 & r_3 & 1 & r_3 & r_6 & r_9 & r_{12} & r_9 \\ r_9 & r_6 & r_3 & 1 & r_3 & r_6 & r_9 & r_{12} \\ r_{12} & r_9 & r_6 & r_3 & 1 & r_3 & r_6 & r_9 \\ r_9 & r_{12} & r_9 & r_6 & r_3 & 1 & r_3 & r_6 \\ r_6 & r_9 & r_{12} & r_9 & r_6 & r_3 & 1 & r_3 \\ r_3 & r_6 & r_9 & r_{12} & r_9 & r_6 & r_3 & 1 \end{bmatrix}.$$

The number of parameters in **W** is now reduced to five: W and the correlations r_3, r_6, r_9, and r_{12} where the indices 3, 6, 9, and 12 refer to time delay in hours between two observations. The sizes of r_3, r_6, r_9, and r_{12} determine the correlations between the corrections. It seems reasonable to let the correlation between the corrections be of the same size as the correlation between the observed differences. These are values that can be estimated easily from time series of observations and short-term forecasts. Autocorrelations of time series for various locations and periods of the year have been calculated, and it was found that the degree of correlation is quite variable. A common feature is that the correlations decay exponentially as a function of time: $r_6 = r_3^2$, $r_9 = r_3^3$, and $r_{12} = r_3^4$. The choice of $r_3 = 0.8$, $r_6 = r_3^2 = 0.6$, $r_9 = r_3^3 = 0.5$, and $r_{12} = r_3^4 = 0.4$ gives acceptable results for stations observing every three hours and for stations observing less than eight times a day, for example, only at 0600, 1200, and 1800 UTC. Whether the overall results can be improved significantly by letting the r_i vary from station to station and through the year is discussed in section 7.

Term **V** is the covariance matrix of the observation error. In our case with only one observation at time t it is reduced to V^2, the variance of the observation error. The weight given to the new observation when updating the corrections is given by the Kalman gain matrix (2.6), which in the case with only one observation at each time step is a vector $\mathbf{k}_t$. Vector $\mathbf{k}_t$ is a function of **W**, V^2, and $\hat{\mathbf{P}}_{t-1|t-1}$, the error covariance matrix of $\hat{\mathbf{x}}_{t-1|t-1}$. Figure 7 gives $k_{1,t}$, the weight given to the new observation when updating $\hat{x}_{1,t}$, as a function of time. The diurnal oscillations illustrate that the new observation is given large weight when updating $\hat{x}_{1,t}$ at 0000 UTC but given small weight when updating $\hat{x}_{1,t}$ at 1200 UTC. The figure gives an idea of how fast $k_{1,t}$ approaches its limit. How fast $\mathbf{k}_t$ reaches the limit values depends on W/V; a large value implies fast convergence, a small value slow convergence. The limit values are independent of the initial value $\hat{\mathbf{P}}_{0|0}$ and depend on W and V only through their relationship W/V. Figure 8 gives the limit values of the daily minimum, mean, and maximum weight given to an observation as a function of W/V. How fast the filter adapts to a new weather situation after a significant change depends on W/V. It is desirable that the filter adapts relatively fast to weather changes. On the other hand, the filter should give stable estimates, where a particular divergent and possibly wrong observation cannot introduce large changes in the corrections from one day to the next. A happy medium was chosen by setting $W/V = 0.06$. In sections 6 and 7 the effects of higher and lower values of W/V are demonstrated.

The task of designing a Kalman filter model includes studies of $\hat{\mathbf{P}}_{t|t}$, the covariance matrix of the errors in the estimated corrections. When updating each time step with one observation, $\hat{\mathbf{P}}_{t|t}$ will also approach a limit. Figure 9 gives $\hat{P}_{1,1,t|t}$ as a function of time. The diurnal oscillations indicate that the error in $\hat{x}_{1,t}$ is lower when updating with an observation at 0000 UTC than at 1200 UTC.

b. Specification of initial conditions

The Kalman filter equations (2.3)–(2.7) recursively update estimates of the unknown state $\hat{\mathbf{x}}_{t|t}$ and their

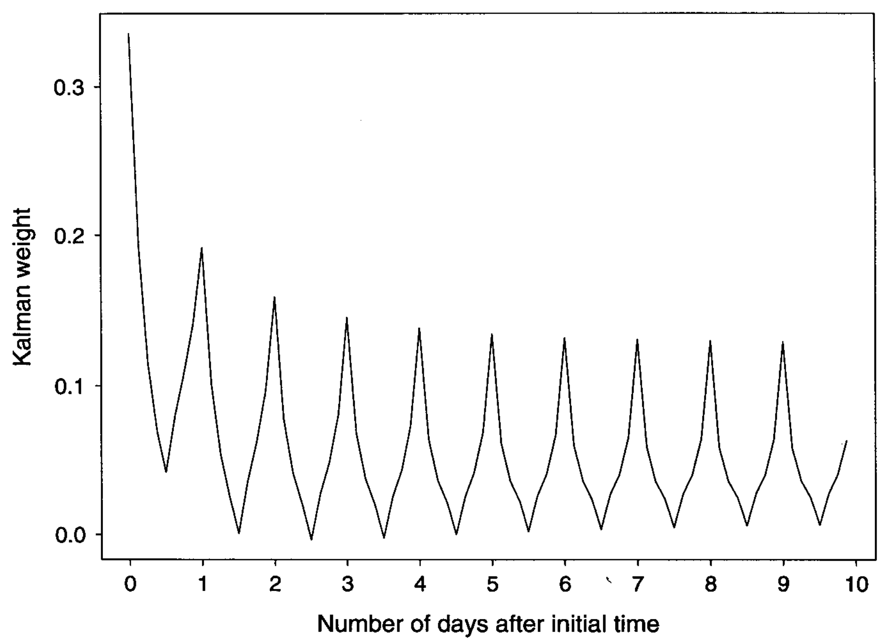

FIG. 7. The Kalman weight $k_{1,t}$ given to the new observation when updating the correction valid at 0000 UTC, as a function of time; $W/V = 0.06$.

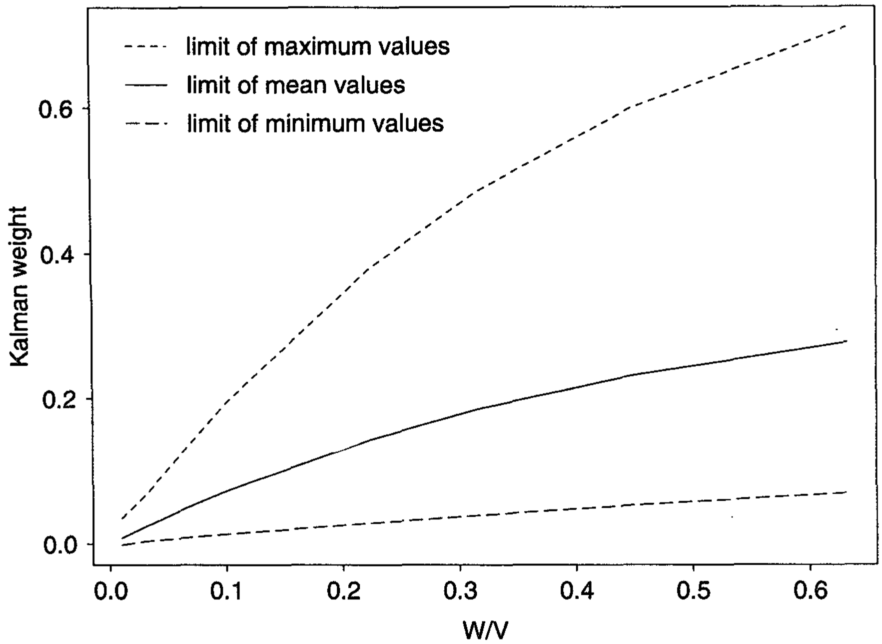

FIG. 8. Limits of the daily maximum (short dashed), mean (full), and minimum (long dashed line) values of the Kalman weight $k_{1,t}$ given to the new observation when updating the correction valid at 0000 UTC, as a function of W/V.

uncertainties $\hat{\mathbf{P}}_{t|t}$. This section has described the specification of the system matrix, the observation matrix, and the covariance matrices. The initial values of the corrections $\hat{\mathbf{x}}_{0|0}$ and their uncertainties $\hat{\mathbf{P}}_{0|0}$ have to be specified as well before running the Kalman filter. In the first time steps, the filter weights are functions of $\mathbf{W}$, V, and $\hat{\mathbf{P}}_{0|0}$, but they will soon converge to values depending only on $\mathbf{W}$ and V. The value of $\hat{\mathbf{x}}_{0|0}$ is unknown, and a large value of $\hat{\mathbf{P}}_{0|0}$ will cause the filter to give more weight to the observations in the first time

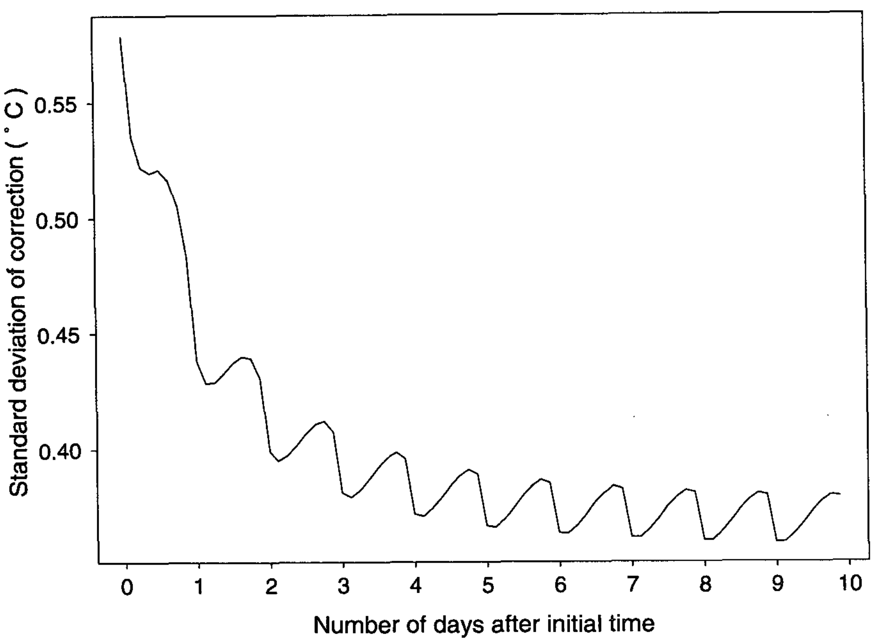

FIG. 9. The standard deviation of the correction valid at 0000 UTC, $(\hat{P}_{1,1,t|t})^{1/2}$, as a function of time; $W/V = 0.06$.

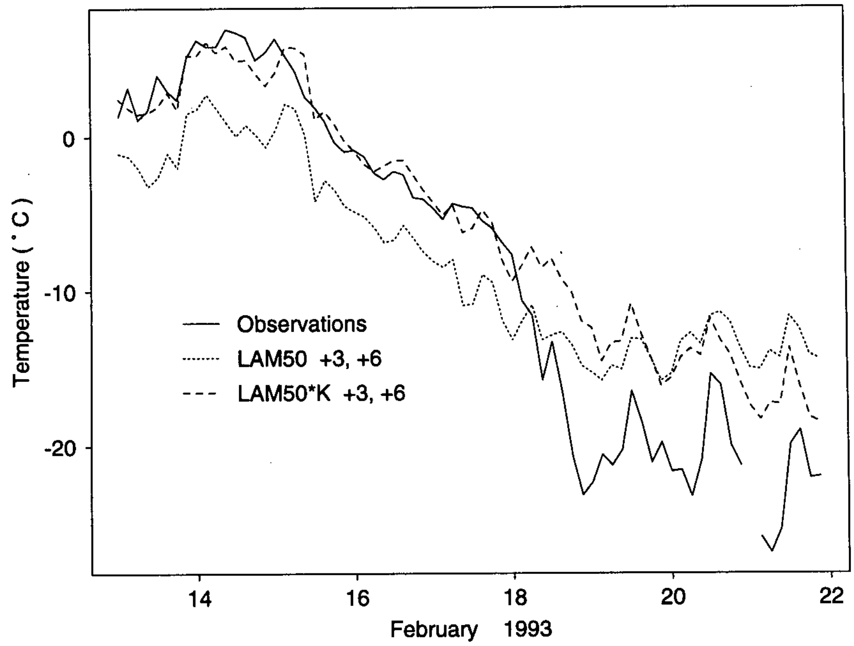

FIG. 10. Temperature observations (full line) and short-term forecasts (+3, +6 h), uncorrected (dotted line) and Kalman filter corrected (dashed line), for Bardufoss from 13 to 21 February 1993.

steps than is given later on. The systematic deviations are set to zero initially. The covariance matrix $\hat{\mathbf{P}}_{0|0}$ is supposed to have a structure identical to $\mathbf{W}$, but with higher variance, for example, $P^2 = 0.5$.

5. Operational implementation and running of the Kalman filter correction procedure

The NWP model LAM50 is run four times a day and also at 0600 and 1800 UTC, even if the number of observations is limited at these times. Surface temperature forecasts (+0, +3, +6, . . . , +42, or +48 h) for about 240 different observation stations are corrected by the Kalman filter model described in section 4, before being presented graphically to the users.

The corrections are updated recursively, following (2.3)–(2.7), at 0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC. As only the corrections calculated at 0000, 0600, 1200, and 1800 UTC are applied to correct temperature forecasts, the operational Kalman filter procedure is only run four times a day: namely, when the 0000, 0600, 1200, and 1800 UTC forecasts are available. At $t \sim 1200$ UTC, for example, the 0600 UTC corrections $\hat{\mathbf{x}}_{t-2|t-2}$ will be updated by the difference between the 06 + 3-h forecast and the observation at 0900 UTC to give $\hat{\mathbf{x}}_{t-1|t-1}$. These corrections are updated once more by the difference between the 0600 UTC + 6-h forecast and the observation at 1200 UTC to give $\hat{\mathbf{x}}_{t|t}$. These corrections are applied to forecasts produced at 1200 UTC. All that need to be saved in a database are the corrections $\hat{\mathbf{x}}_{t|t}$ and their uncertainties $\hat{\mathbf{P}}_{t|t}$ for each observing station.

The Kalman filter deals with missing observations in a satisfactory manner. Setting the variance of the observation $V_t = \infty$ in (2.6) gives $K_t = 0$, which implies that the updated corrections are equal to the previous ones: $\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} = \hat{\mathbf{x}}_{t-1|t-1}$. However, the uncertainties in the corrections are increased: $\hat{\mathbf{P}}_{t|t} = \hat{\mathbf{P}}_{t|t-1} = \hat{\mathbf{P}}_{t-1|t-1} + \mathbf{W}_t$. This implies that the same Kalman filter equations can be applied to all observing stations, the ones observing every three hours and the ones observing less than eight times a day.

6. Examples of performance

In this section, the behavior of the Kalman filter model presented in the previous sections will be illustrated by some figures giving time series plots of temperature observations and short-term forecasts, both uncorrected and corrected. Various situations when the filter works well and situations that are difficult to correct will be highlighted.

Figure 10 gives results for Bardufoss during the period 13–21 February 1993. From 13–18 February the short-term (+3 and +6 h) forecasts are systematically about 5°C below the observations. The Kalman filter works well and gives corrected temperature forecasts that are very close to the observed temperatures. On 18 February the weather changes, giving very low temperatures during the night, more than 5°C lower than forecasted. The corrections for the 18 February are still positive, giving corrected temperatures that are 10°C too warm! However, the forecasts continue being too high, and the corrections are reduced to zero after one day. From the 20 February, the corrections become negative, bringing the corrected forecasts closer to the observations. In this situation the response of the filter is clearly too slow. However, it will be shown later on that there are situations where such a slow response is advantageous.

The filter's ability to correct for systematic diurnal varying deviations between forecasts and observations is demonstrated with results for Færder, a lighthouse in the Oslofjord. The observed temperatures are influenced by the sea temperature, implying small diurnal variation as compared to the NWP forecasts. The time series plot for Færder from 5 to 15 April 1993 (Fig. 11) illustrates that the corrected temperatures are much closer to the observed than the uncorrected. Notice also how small the corrections are on 8 April after a period with very unsystematic deviations between forecasts and observations.

Other locations will typically have too small diurnal amplitudes in the forecasted temperatures. Nesbyen, situated in a valley in the southern part of Norway, is observing four times a day only, at 0600, 1200, 1500, and 1800 UTC. The data presented in Fig. 12 from 9 to 15 May 1993 show an increased diurnal amplitude of the corrected forecasts. The filter also gives reasonable corrections at 0000, 0300, 0900, and 2100 UTC, times with no observations. Notice how the filter behaves when the stable pattern is broken by one different day, 13 May. The corrections are reduced, but still positive, implying good results on the following day when the stable pattern is resumed.

As pointed out earlier, the relationship W/V determines how sensitive the filter is to spurious changes in the deviations between forecasts and observations and also how fast the filter adapts to changes. In Figs. 10–12, $W/V = 0.06$ (the value applied operationally). Higher values of W/V will give a filter that is more sensitive to the latest observations, while lower values of W/V will give a more stable filter. As an example, results for Bardufoss from 16–21 February 1993 with different values of W/V are presented in Fig. 13. Here, $W/V = 0.16$ implies a very fast adaptation to the observations after the weather change, while the corrections implied with $W/V = 0.01$ are very conservative.

7. Summary statistics

In this section, the quality of the Kalman filter corrected forecasts for different values of the covariance matrices will be studied. Monthly, seasonal, and yearly values of bias, standard deviations, and mean absolute errors of the forecasts have been calculated for the stations examined in section 3. Some typical results will be presented.

The correlations r_3, r_6, r_9, and r_{12} in the covariance matrix **W** determine the degree of correlation between the eight corrections. Here, $r_3 = r_6 = r_9 = r_{12} = 1$ implies that the eight corrections are equal at each time step, $r_3 = r_6 = r_9 = r_{12} = 0$ implies that they are independent, while the operational chosen values $r_3 = 0.8$, $r_6 = r_3^2 = 0.6$, $r_9 = r_3^3 = 0.5$, and $r_{12} = r_3^4 = 0.4$ give correlated but diurnally varying corrections. The results obtained with $r_3 = r_6 = r_9 = r_{12} = 1$ and the operational values

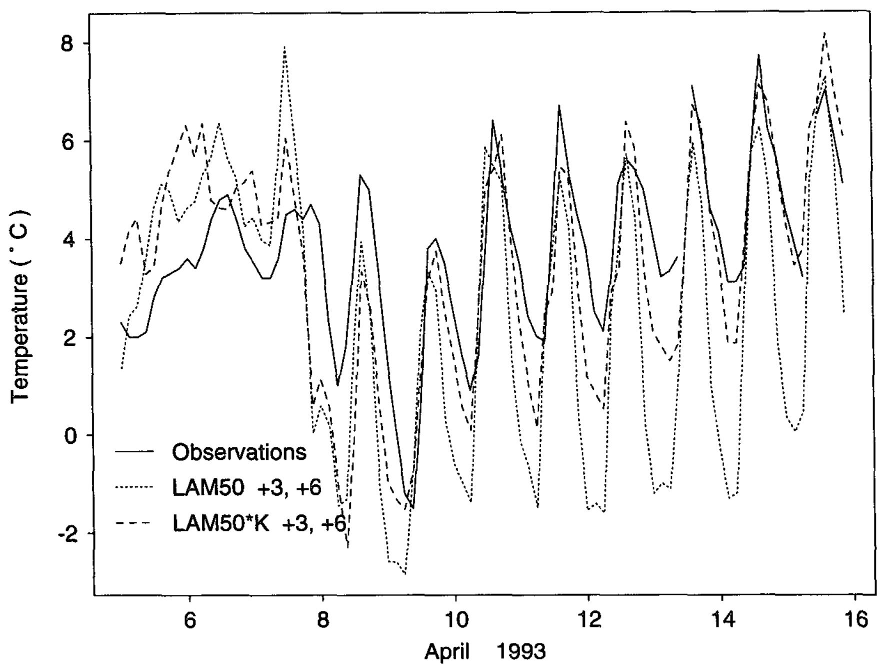

FIG. 11. As in Fig. 10 but for Færder from 5 to 15 April 1993.

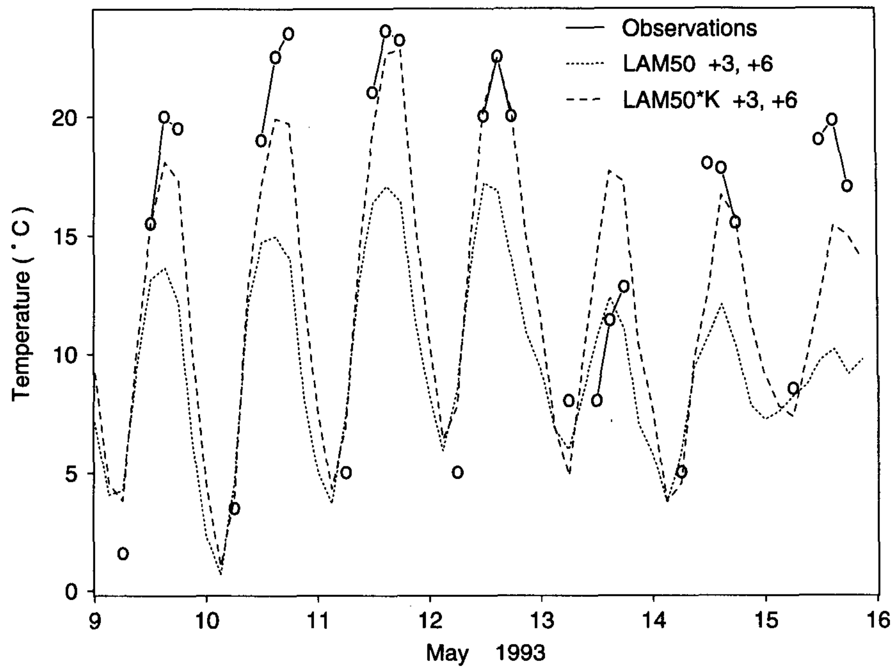

FIG. 12. As in Fig. 10 but for Nesbyen from 9 to 15 May 1993.

will be presented. When studying the sensitivity of the results to the specification of r_i, only short-term forecasts have been evaluated, and the summary statistics have been calculated as a function of valid time. The +3-h forecasts issued at 0000, 0600, 1200, and 1800 UTC are validated at 0300, 0900, 1500, and 1800 UTC,

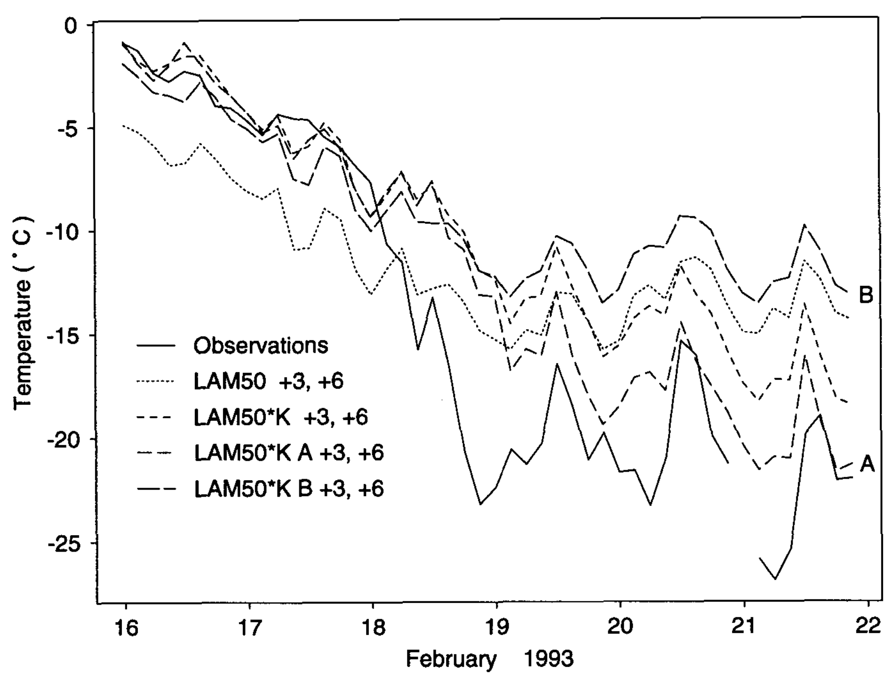

FIG. 13. As in Fig. 10 but for Bardufoss from 16 to 21 February 1993 and with three sets of corrected temperature forecasts; (a) $W/V = 0.16$ (medium dashed line), (b) $W/V = 0.01$ (long dashed line), and (c) as before $W/V = 0.06$ (short dashed line).

respectively, while +6-h forecasts are validated at 0600, 1200, 1800, and 0000 UTC. The results were more sensitive to the values of r_i when there was a systematic diurnal cycle in the forecast errors. Figure 14 presents results at Færder 1993, with the left panel representing the winter months November to March and the right panel representing the summer months May to September, which have significant diurnal cycle (see Fig. 5). The biases of the corrected forecasts were very close to zero when $r_3 = 0.8$, while $r_3 = 1$ only reduced the diurnal mean value of the biases to zero, and the diurnal cycle of the biases remained. The standard deviations of the errors in the short-term corrected forecasts were slightly reduced, both with $r_3 = 0.8$ and $r_3 = 1$. It seemed to be a general feature that the results were significantly better with $r_3 = 0.8$ when there was a systematic diurnal cycle in the forecast errors, while $r_3 = 1$ gave results that were slightly better when there was not.

The relationship W/V determines how much weight is to be given to the new observation when updating the Kalman filter corrections. Three different values of W/V have been evaluated: $W/V = 0.16$, $W/V = 0.06$, and $W/V = 0.01$. As the sensitivity to the specification of W/V depends on forecast projection time, the summary statistics have been calculated as a function of forecast projection. All values are daily mean values, for example, the mean values of +6-h forecasts will include 00 + 6-, 06 + 6-, 12 + 6-, and 18 + 6-h forecasts. At Tafjord, the biases of the LAM50 forecasts were close to −5°C for all forecast projections, and the biases of the corrected forecasts were all close to zero (Fig. 15). When the biases of the LAM50 forecasts increase or decrease as a function of forecast projection time, the biases of the corrected forecasts were reduced to zero only for the short-term forecasts (see the results in Fig. 15 for Færder where the biases of the LAM50 forecasts increase). The results for Tafjord and Færder indicate that the sensitivity to the specification of W/V was low with respect to bias reduction, but the monthly biases of values corrected with $W/V = 0.01$ were in many cases found to be significantly different from zero, as well as for short-term forecasts. It is desirable that the corrected forecasts are unbiased, while at the same time, the standard deviations of the differences between corrected forecasts and observations should be as small as possible. The Kalman filter with $W/V = 0.16$ gave corrections with large variability; the standard deviations of the short-term forecasts were reduced, but such a high value of W/V often implied increased standard deviations of the long-term forecasts (see Fig. 15). The conservative value of $W/V = 0.01$ generally implied less reductions of the standard deviations of the short-term forecasts but larger reductions or only minor increases of the standard deviations of the long-term forecasts.

For the operationally running Kalman filter, a value in between has been chosen: $W/V = 0.06$. This value gave corrected forecasts that were approximately unbiased, and the standard deviations of the differences remained essentially unchanged. This has been verified by summary statistics calculated for about one-half of the observing stations, the ones with data stored in a database and that are observing regularly at least at 0000, 0600, 1200, and 1800 UTC. The results of the winter months November to March 1993 and 1994 were quite close and have been gathered in the left panel of Fig. 16. The results of the summer months May to September 1993 and 1994 are presented in the right panel. The mean absolute errors of the uncorrected forecasts were higher in winter and increased slightly as a function of forecast projection time. The reductions of the mean absolute errors implied by the correction procedure were larger in the summer, about 1°C for all projections. The reductions in the mean absolute errors reflect the bias reduction at each observing station. To illustrate how close to zero the monthly biases of the corrected forecasts at each observing station were, *averages of the absolute values* of the monthly biases at each observing station are also presented in Fig. 16. The bias reduction was most successful in summer and for short-term forecasts. The mean values of monthly standard deviations were slightly reduced in summer but increased slightly for the long-term forecasts in winter.

8. Discussion and ideas for further improvements

It has been shown that there are systematic deviations between surface temperature forecasts produced by the LAM50 NWP model and the corresponding observations that vary dramatically with respect to time of day, month of year, and location. Accordingly, a Kalman filter model has been defined to correct the forecasts, allowing for diurnally varying corrections. Eight corrections, valid at 0000, 0300, 0600, 0900, 1200, 1500, 1800, and 2100 UTC, are calculated simultaneously.

Among the advantages of the Kalman filter as compared to traditional statistical methods are that the corrections are continuously updated with the latest observations, allowing it to adapt when there are major changes in the NWP model. Another advantage is its simplicity. With the same set of statistical parameters, it worked at all observing stations throughout the year and caused the monthly bias in the forecasts at each location to be reduced to values close to zero. The monthly standard deviations of the differences remained essentially unchanged.

The assumptions in section 4 leading to $\mathbf{W}$ and $\mathbf{V}$ may be questioned. The sensitivity of the final results to the specification of the correlations, as illustrated in the previous section, indicates which improvements are to be expected with $\mathbf{W}$ and $\mathbf{V}$ depending on location and season. Another improvement would be to update different set of corrections for different forecast projection times, with W/V decreasing with forecast pro-

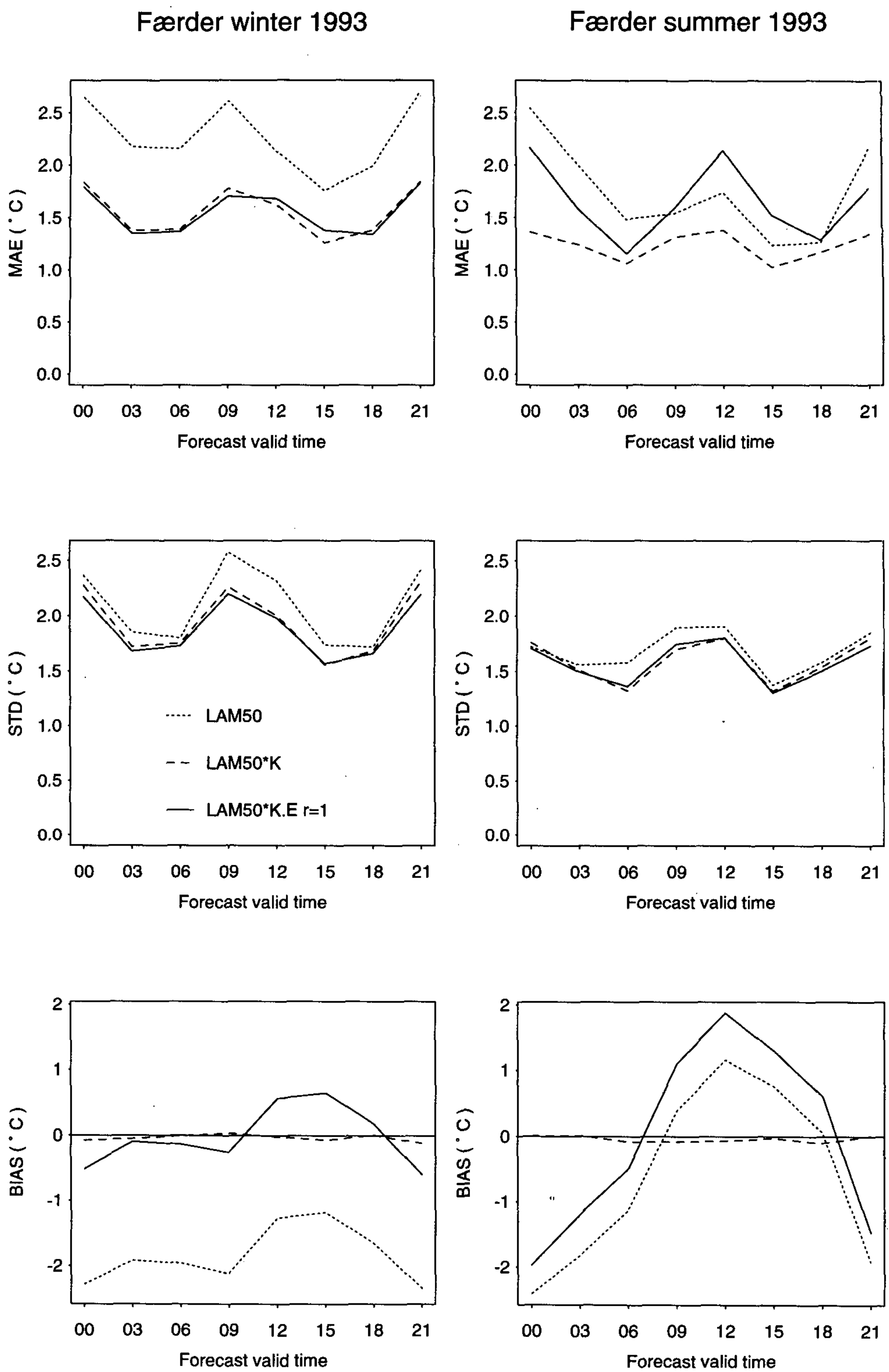

FIG. 14. Summary statistics for Færder 1993 presented as a function of forecast valid time. The left panel gives winter (November to March) results and the right panel summer (May to September) values. Mean absolute error (MAE) (upper panel), standard deviations (STD) (mid panel), and bias (lower panel) are calculated for differences between short-term (+3, +6 h) temperature forecasts from LAM50 and observations, direct model output (dotted lines), and forecasts corrected with the operational Kalman filter procedure with $r_3 = 0.8$ and $W/V = 0.06$ (dashed lines) and with $r_3 = 1.0$ and $W/V = 0.06$ (full lines).

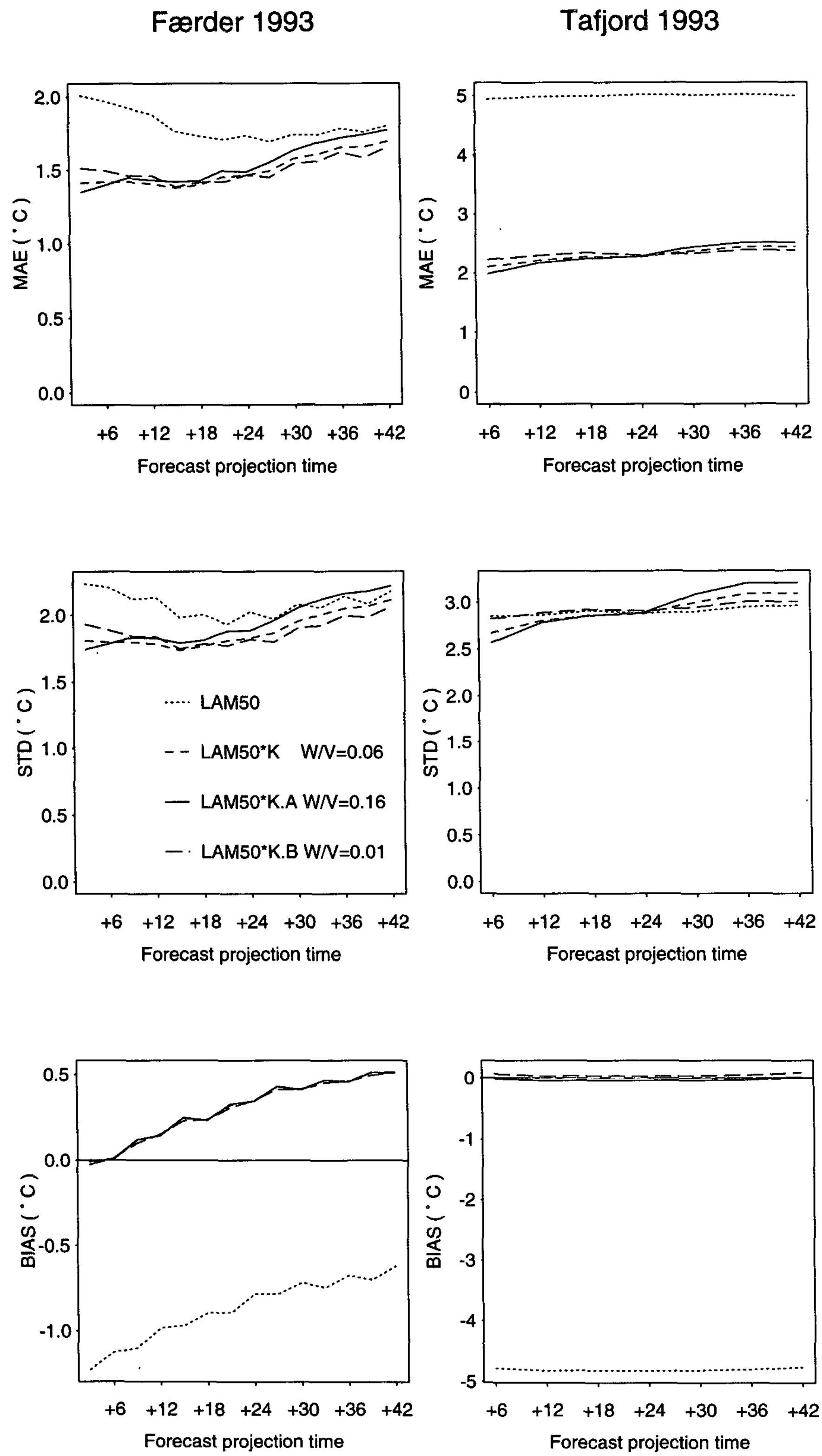

FIG. 15. Summary statistics for Færder (left panel) and Tafjord 1993 (right panel) presented as a function of forecast projection. MAE (upper panel), STD (mid panel), and bias (lower panel) are calculated for differences between temperature forecasts from LAM50 and observations, direct model output (dotted lines), and forecasts corrected with the operational Kalman filter procedure with $r_3 = 0.8$ and $W/V = 0.06$ (dashed lines), with $r_3 = 0.8$ and $W/V = 0.16$ (full lines), and with $r_3 = 0.8$ and $W/V = 0.01$ (long dashed lines).

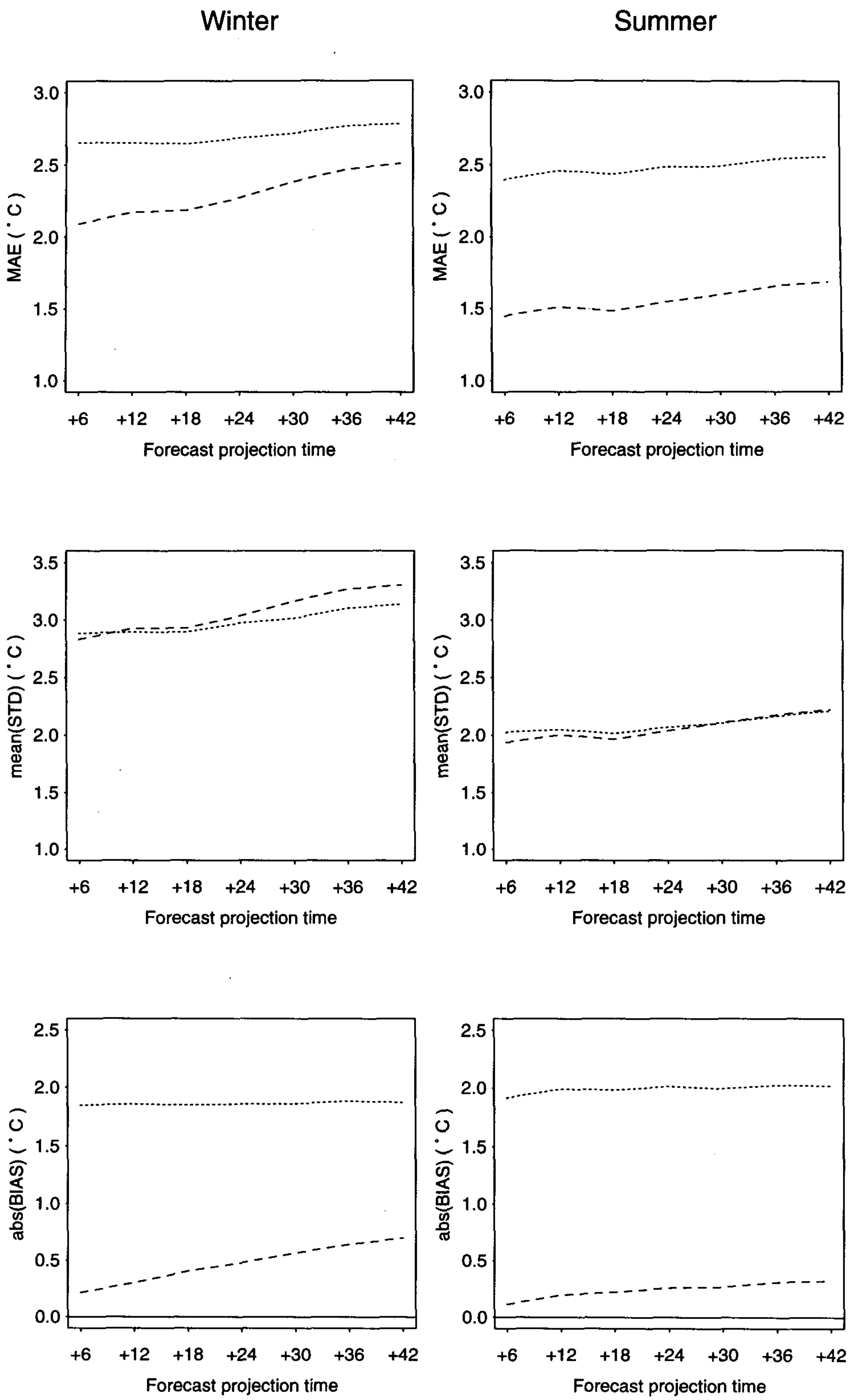

FIG. 16. Summary statistics calculated for about 120 observing stations for the winter months from November to March 1993 and 1994 (left panel) and the summer months from May to September 1993 and 1994 (right panel) presented as a function of forecast projection. MAE (upper panel), mean STD (mid panel), and mean of absolute values of monthly biases [abs(BIAS)] (lower panel) are calculated for differences between temperature forecasts from LAM50 and observations, direct model output (dotted lines), and forecasts corrected with the operational Kalman filter procedure with $r_3 = 0.8$ and $W/V = 0.06$ (dashed lines).

jection time. In both cases, time series of temperature forecasts and observations at each station would have to be studied. It is not clear that this is worthwhile.

The main weakness of the correction procedure is the behavior at sudden weather changes. The correction of future values is only based on previous differences between temperature observations and forecasts. The model does not use any knowledge of the weather dependency of the systematic deviations. Persson (1991) demonstrates improved results on some occasions, with a Kalman filter model assuming a linear dependency between the errors of the forecasts and the temperature forecasts at the surface and 850 hPa. It is possible, but not as simple, to make a corresponding extension of our Kalman filter model. However, there are other problems with such a formulation, because the dependencies between the errors of the temperature forecasts and other prognostic values are not always linear. Moreover, the importance of different predictors is strongly dependent on geographical position: in some places the wind is more important, in other places the stability or cloud conditions are most important. The dependencies, if any, are also varying in time.

Persson (1991) has another interesting idea, that of manipulation of the variances W^2/V^2 through external interference. The change of the statistical parameters could also be done automatically, in accordance with weather changes identified in the NWP forecasts.

The problem with the behavior of the filter at sudden weather changes is not a trivial one and the need for a solution becomes even more pronounced when correcting forecasts valid more than two days ahead. If the weather giving rise to the most extreme errors in the temperature forecasts can be identified, it is possible to calculate and apply different sets of corrections for different weather. The weather identification has to be dynamic, since it is known that the dependency of the error on weather parameters varies both in time and space.

Acknowledgments. This study is supported by the Norwegian Research Council, project number 100600/410. Nils Lid Hjort, Department of Mathematics and Statistics, University of Oslo, was my Ph.D. supervisor. Erik Bølviken, Department of Mathematics and Statistics, University of Oslo, has given valuable support through discussions in the early stage of the work. Audun D. Christoffersen and Anstein Foss have assisted the implementation of the operational Kalman filter correction procedures, which ran smoothly. I am also grateful to Magne Aldrin, Norwegian Computing Centre, Oslo, Jan Erik Haugen and Thor Erik Nordeng, DNMI, and Anders Persson, ECMWF, for their comments on the paper and to the two reviewers for their careful reading of the manuscript and their comments that helped me make things clearer.

REFERENCES

Bohlin, T., 1976: Four cases of identification of changing systems. *System Identification: Advances and Case Studies.* R. Mehra and D. G. Laniotis, Eds., Academic Press.

Bratseth, A. M., 1986: Statistical interpolation by means of successive corrections. *Tellus,* **38A,** 439–447.

Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977: Maximum likelihood from incomplete data via the EM Algorithm. *J. Roy. Statist. Soc. Ser. B,* **39,** 1–38.

Gelb, A., 1974: *Applied Optimal Estimation.* MIT Press, 374 pp.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

Grønås, S., and K. H. Midtbø, 1987: Operational multivariate analysis by successive corrections. *Proc. WMO/IUGG NWP Symp.,* Tokyo, Japan, Meteor. Soc. Japan, 61–74.

Harrison, P. J., and C. F. Stevens, 1976: Bayesian forecasting. *J. Roy. Statist. Soc. Ser. B,* **38,** 205–248.

Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *Trans. ASME, Ser. D,* **82,** 35–45.

——, 1963: New methods of Wiener filtering theory. *Proceedings First Symposium on Engineering Appns. of Random Functions Theory and Probability,* J. L. Bogdanoff and F. Kozin, Eds., Wiley.

——, and R. S. Bucy, 1961: New results in linear filtering and prediction problems. *Trans. ASME, Ser. D,* **83,** 95–108.

Kilpinen, J., 1992: The application of Kalman filter in statistical interpretation of numerical weather forecasts. *Proc. 12th Conf. on Probability and Statistics in Atmospheric Sciences,* Toronto, ON, Canada, Amer. Meteor. Soc., 11–16.

O'Hagan, A., 1978: Curve fitting and optimal design for prediction. *J. Roy. Statist. Soc. Ser. B,* **40,** 1–42.

Persson, A., 1991: Kalmanfiltering—a new approach to adaptive statistical interpretation of numerical meteorological forecasts. *Lectures and Papers Presented at the WMO Training Workshop on the Interpretation of NWP Products in Terms of Local Weather Phenomena and Their Verification,* Wageningen, the Netherlands, WMO, XX-27–XX-32.

Priestley, M. B., 1981: *Spectral Analysis and Time Series.* Academic Press, 807–815.

Ross, G., 1989: Model output statistics using an updatable scheme. *Proc. 11th Conf. on Probability and Statistics in Atmospheric Sciences,* Monterey, CA, Amer. Meteor. Soc., 93–97.

Simonsen, C., 1991: Self adaptive model output statistics based on Kalman filtering. *Lectures and Papers Presented at the WMO Training Workshop on the Interpretation of NWP Products in Terms of Local Weather Phenomena and Their Verification,* Wageningen, the Netherlands, WMO, XX-33–XX-37.

Young, P. C., 1974: Recursive approaches to time series analysis. *Bull. Inst. Math. Appl.,* **10,** 209–224.