Indian J. pure appl. Math., **31**(8) : 921-932, August 2000

FORECAST SURFACE TEMPERATURE CORRECTION BY ESTIMATION THEORY

KAUSTUBHA BHATTACHARYA

NCMRWF, Mausam Bhavan Complex, Lodhi Road, New Delhi 110 003, *India*

(*Received* 1 *October* 1998; *after revision* 11 *March* 1999; *accepted* 25 *August* 1999)

Surface temperatures are difficult to forecast using a General Circulation Model because of various imperfections in the latter. In this paper, a new post-processing method has been reported, whereby 3rd day forecasts for surface temperature at Delhi (at 1.36m) are updated using sequential estimation theory. The method, with a Kalman filter, is recursive because only the latest correction (which contains the influence of past corrections in a weighted manner) is used to obtain future corrections. The difference between 3-hourly synoptic data for December 1995 from Safdarjang, New Delhi and the corresponding short term forecasts are inserted as input to "observations" for the Kalman filter scheme. The results are displayed for three different periods: for the beginning, middle and end of the month. Daily maximum and minimum temperatures before and after correction are also shown, along with observed values. The scope for extension of the work is discussed.

Key Words : Kalman Filter; Forecast; Temperature Correction

1. INTRODUCTION

Numerical Weather Prediction (NWP) models are known to have systematic errors from different sources. The finite resolution of such models presents difficulty in the representation of orography, giving rise to difference in station elevations. This problem is felt even for models with small grid size (Cattani, 1994)[1]. The other known sources of major errors in forecast fields include inadequate parameterizations of the boundary layer and other physical processes, and incorrect representations of ground conditions. These give rise to deficiencies in forecast fields, with the surface forecast being particularly erroneous.

While the understanding and reduction of such systematic errors at source is a long term scientific goal, there is the more immediate practical need of post-processing the "Direct Model Output" forecasts before issuing it to the public. The conventional post-processing (or statistical interpretation) methods, such as, Model Output Statistics (MOS) or Perfect Prog Method (PPM) utilize a two-step procedure. The steps are: (1) evaluate the statistical link between the predictor (upper air fields) and the predict and (surface field) — these fields can be obtained either from observation (for PPM) or forecast (for MOS); and (2) use these pre-determined statistical links to prepare the actual forecast. The conventional methods, however, have two serious drawbacks: (1) a long time series of data are needed to establish stable statistical relationships; and (2) if it is necessary to find new links then one has to construct the new time series of data all over again. This is the case whenever there is a change in the forecast model, or data assimilation scheme, and such activities are normal for any progressive NWP centre.

In this paper, a recursive approach for the adaptation of the abovementioned statistical links — the Kalman filter — is described. In this method, the statistical relationship between the predictor and the predictand ("Kalman gain") is continuously and automatically updated with the help of the forecast error (dynamic) and observation error (static) covariances. The principle behind this new approach has been explained by the work performed at the NCMRWF to correct the model forecast for 1.36m temperature at Delhi for the month of December 1995.

Historically, in the context of data assimilation the Kalman filter was preceded by the Wiener filter (Wiener, 1956)[14]. Ignoring dynamics completely, Wiener's method used the knowledge of a system's complete past history, together with the Wiener-Hopf (filter) equations to estimate the system's present and future. Unfortunately, the Wiener filter was unwieldy and not suitable for use in real time; therefore, it was not widely used.

Since the diurnal cycle is strong in the tropics (Hastenrath, 1991)[7], and this phenomenon is simulated in operational surface weather forecasts, it is necessary to operate the Kalman filter (or any other post processing method) with a time resolution that is good enough to resolve the diurnal cycle. But, how frequently this can be done depends on the data availability. As model outputs can be utilized even for every time step, the best time resolution possible with the India Meteorological Department's (IMD) synoptic network is 3 hours as some stations in the network (including Delhi) take observations at that interval.

2. METHOD

The Kalman filter was developed in the 1960's for estimation of stochastic variables that arise from randomly perturbed differential or difference equations. It has been widely used in aerospace problems that involve optimal control and signal processing. For instance, given the current position and velocity of a spacecraft, the Kalman filter can optimally estimate its future position and velocity. Its use in meteorology was pioneered by Ghil *et al.* (1981)[4] for data assimilation. Today, the Kalman filter is utilized in several other branches of meteorology, such as NWP applications (Persson, 1991)[12] and wind estimation from tracer measurements (Daley, 1996).[2] It is also used in oceanography (Ghil and Malannotte-Rizzoli, 1991)[6].

The Kalman filter theory is presented here briefly. The details are available in Gelb (1974)[3], Jazwinski (1970)[9] or Kailath (1968)[10].

Two basic assumptioms are needed to explain the Kalman filter. The first concerns the evolution of the true state of the system

$$W_k^t = F_{k|k-1} W_{k-1}^t + b_{k-1}^t, \qquad \text{... (2.1)}$$

where W_k^t is the (unobservable) true model state at time step k, $F_{k|k-1}$ is a matrix that describes the dynamics of the system for transition from $k-1$ to k and b_k^t *represents a random error (system noise) during change in the state. This is commonly assumed to be a Gaussian white noise process, i.e., errors are not correlated in space or time. This assumption holds true to a large extent for current NWP models in mid latitudes (Ghil, 1989)*[5]. As we shall see, in our case the transition matrix is an identity matrix - so this assumption is not a problem for us.

The second assumption relates the observation W_k^0 at time step k to the true model state W_k^t by

$$W_k^0 = H_k W_k^t + b_k^0 \qquad \text{... (2.2)}$$

where H_k is the observation operator at time step k. It is a matrix that describes the correspondence between w_k^t and W_k^0, which is sometimes difficult to construct in practice, and b_k^0 is again assumed to be a (Gaussian) white noise random process (observation) noise. While this last assumption is true for many observations, it is not true for all kinds of observations. However, these two assumptions about noise processes are not crucial to the subsequent results (Ghil, 1989).

One more assumption is needed for presentation of the Kalman filter, and it is that the system noise b_k^t and observation noise b_k^0 are uncorrelated with each other.

There might be cases where the assumptions about observation errors may be violated, as for instance in temperatures derived from satellite radiance measurements (which may be correlated in space and time) or in geopotential heights derived from radiosonde measurements. In addition, as often is the case with satellite measurements, if short-range model forecasts are used to obtain temperatures from radiances, then the last assumption will also be violated. Fortunately all these cases can be easily dealt with by minor modifications of the Kalman filter equations (Gelb, 1974)[3].

The algorithm for utilising the Kalman filter theory to recursively estimate the state vector $W_{k|k}$ is performed in two steps. $W_{k|k}$ is the state vector at time step k where the observation at time step k has been utilized. The first step is the prediction step, where the transition matrix $F_{k|k-1}$ is used to advance the model state from time step $k-1$ to time step k, assuming the process was observed upto time step $k-1$ and an optimal estimate of the state upto $k-1$ is available. The equation used in this step (assumed to be imperfect) is

$$W_{k|k-1} = F_{k|k-1} W_{k-1|k-1}. \qquad \text{... (2.3)}$$

If $P_{k|k-1}$ is the covariance matrix of the error in the state vector, then

$$P_{k|k-1} = F_{k|k-1} P_{k-1|k-1} F_{k|k-1}^T + Q_{k-1} \qquad \text{... (2.4)}$$

where Q_k estimates the covariance matrix of the system's noise, i.e., the "model error" and the superscript 'T' refers to the transpose of the matrix. We have

$$E b_k^t (b_l^t)^T = Q_k \, \delta_{kl}, \qquad \text{... (2.5)}$$

where E is the expectation operator representing the mean or average of a set of infinite number of measurements, and δ_{kl} is the Kronecker delta.

The next step is the update step. It uses the observation available at step k to update the previous estimate of the state $W_{k|k-1}$ by forming a linear combination of the previous estimate and the prediction error

$$W_{k|k} = W_{k|k-1} + K_k (W_k^0 - H_k W_{k|k-1}), \qquad \text{... (2.6)}$$

where

$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \qquad \text{... (2.7)}$$

is the "Kalman gain" matrix, and R_k is the covariance matrix of observation noise (observation error), i.e.,

$$E\, b_k^0 (b_l^0)^T = R_k\, \delta_{kl}. \quad \text{... (2.8)}$$

The covariance matrix of the error in $W_{k|k}$ is now given by

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} \quad \text{... (2.9)}$$

K_k, the "Kalman gain" matrix (a term borrowed from engineering literature) is more commonly known as the "weight" matrix in meteorology. In fact, all data assimilation systems follow the two eqs. (2.3) and (2.6), the difference between them being in methods of computing K_k. The success of the Kalman filter (as compared to other assimilation methods) lies in the fact that the error covariance matrix ("forecast error") $P_{k|k-1}$ is updated at each step, and so is K_k. The column vectors H_k, K_k and R_k have in our case 8, 8 and 8 elements respectively, while the matrix $P_{k|k}$ is a 8 × 8 matrix.

The Kalman fitler algorithm is presented in Table I using the equations presented above. Similar tables for Kalman, Wiener and adaptive filter algorithms can be found in Prasad, *et al.* (1995)[13].

TABLE I : The Kalman filter algorithm using recursive predictor-corrector approach

Initial conditions

$$W_{0|0} = O$$

$$P_{0|0} = E\,[W_{0|0}\, W_{0|0}^T]$$

Compute for n = 1, 2, 3, ...

$$W_{k|k-1} = F_{k|k-1}\, W_{k-1|k-1}$$

$$P_{k|k-1} = F_{k|k-1}\, P_{k-1|k-1}\, F_{k|k-1}^T + Q_{k-1}$$

$$K_k = P_{k|k-1}\, H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1}$$

$$W_{k|k} = W_{k|k-1} + K_k (W_k^0 - H_k\, W_{k|k-1})$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1}.$$

In the present work, the T80 Direct Model Output (DMO) of the 3rd day operational NCMRWF surface air temperature forecast for Delhi at 1.36m was corrected every 3 hours (i.e., at 00Z, 03Z, 06Z, 09Z, 12Z, 15Z, 18Z, 21Z). These corrections were performed for every day of December, 1995. The synoptic observation (conventionally at 1.36m) are taken at the same times at Safdarjung Airport, New Delhi Observatory. Short term (6 h) forecasts at 1.36m were available at the same time from the archived products of the operational data assimilation scheme at NCMRWF, for the main synoptic hours (i.e., 00Z, 06Z, 12Z, 18Z). For the intermediate hours (03Z, 09Z, 15Z, 21Z) the short term forecasts at 1.36m were generated by running the T80 forecast model for 3 hours starting from 00Z, 06Z, 12Z, 18Z initial conditions. The difference between synoptic observation and DMO of the 3rd day of daily 5-day operational forecast valid at the same time was taken to be the "state vector", while the difference between the synoptic observation and short term forecast valid at the same time was taken to be the "observation". This philosophy is entirely consistent with that behind the use of observations in data assimilation, where the difference between the observation and background information ("guess") is used to update field values.

For the Kalman filter the transition and observation matrices $F_{k|k}$ and H_k had to be specified, along with the covariance matrices for the system noise, and observation noise (Q_k and R_k). The transition matrix was taken to be an identity matrix because in the absence of any quantitative knowledge about the reason for change in the DMO forecast with time, it was assumed that only a random process was responsible for the change. This is a valid assumption because here we are concerned with 3-day forecasts. The observation operator was a matrix whose elements were such that for updating the state vector for a given time of the day, the filter uses the observation for that time only. In addition, the starting values of the state vector and the corresponding error covariance matrix need to be specified, but here the exact values are not so crucical for a linear calculation due to a rapid convergence to asymptotic values (Ghil and Malanotte-Rizzoli, 1991)[6] which is our case. On the other hand, the choice of the matrices describing the system and observation noises are crucial (Homleid, 1995)[8]. The ratio of the system and observation variances determine the speed of response of the Kalman filter to changing weather situations. Also, the relative values of the off-diagonal elements of the system noise matrix determine the correlations between model errors at different valid times of the day. Following Homleid (1995)[8], we assumed that the correlations decay exponentially with time. The values of the diagonal elements of the system noise matrix as well as the initial values error covariance matrix for the state vector were estimated from autocorrelations of the long time series of observation minus DMO forecast differences. This time-series was created from one year of differences calculated from the NCMRWF archives.

The above discussion concerns the inputs to the Kalman filter scheme. For the computer program of the scheme itself, subroutine KALMN from the IMSL statistical library was used after it was tested and found to work satisfactorily.

3. RESULTS

After some trials, all elements, of the observation noise matrix were arbitrarily assigned the value of 25.0, while all elements of the initial state vector error covariance matrix were assigned the common values of 5.0. The initial values of the elements of the state vector were taken to be the difference between observation and DMO forecast at the given time. In formulating the model error covariance matrix, it was postulated that the correlations between corrections at different times were exponentially related; this was based on Homleid's (1995)[8] finding from studies of the decay of autocorrelation of a time series is of observed differences. However, the assumed rate of this decay (assumed to be uniform) for realistic results was stronger than that reported by Homleid (1995)[8]. In other words, we need to assume weaker correlated model errors between successive forecast times on both sides of the main diagonal of the model error matrix.

Fig. 1 shows the 72h forecast, the corresponding verifying observation and the corrected forecast for days 3, 4 and 5 of December 1995. The forecasts started three days earlier at 00Z, and the diurnal oscillation is clear in all the three components. Fig. 2 shows the same set of plots for days 13, 14, 15 and Fig. 3 shows the same set for days 23, 24, 25. The three figures show the situation at the beginning, middle and the end of the month. We find that the corrected forecast generally follows the trend of the observations, and is usually close to them too. However, on a

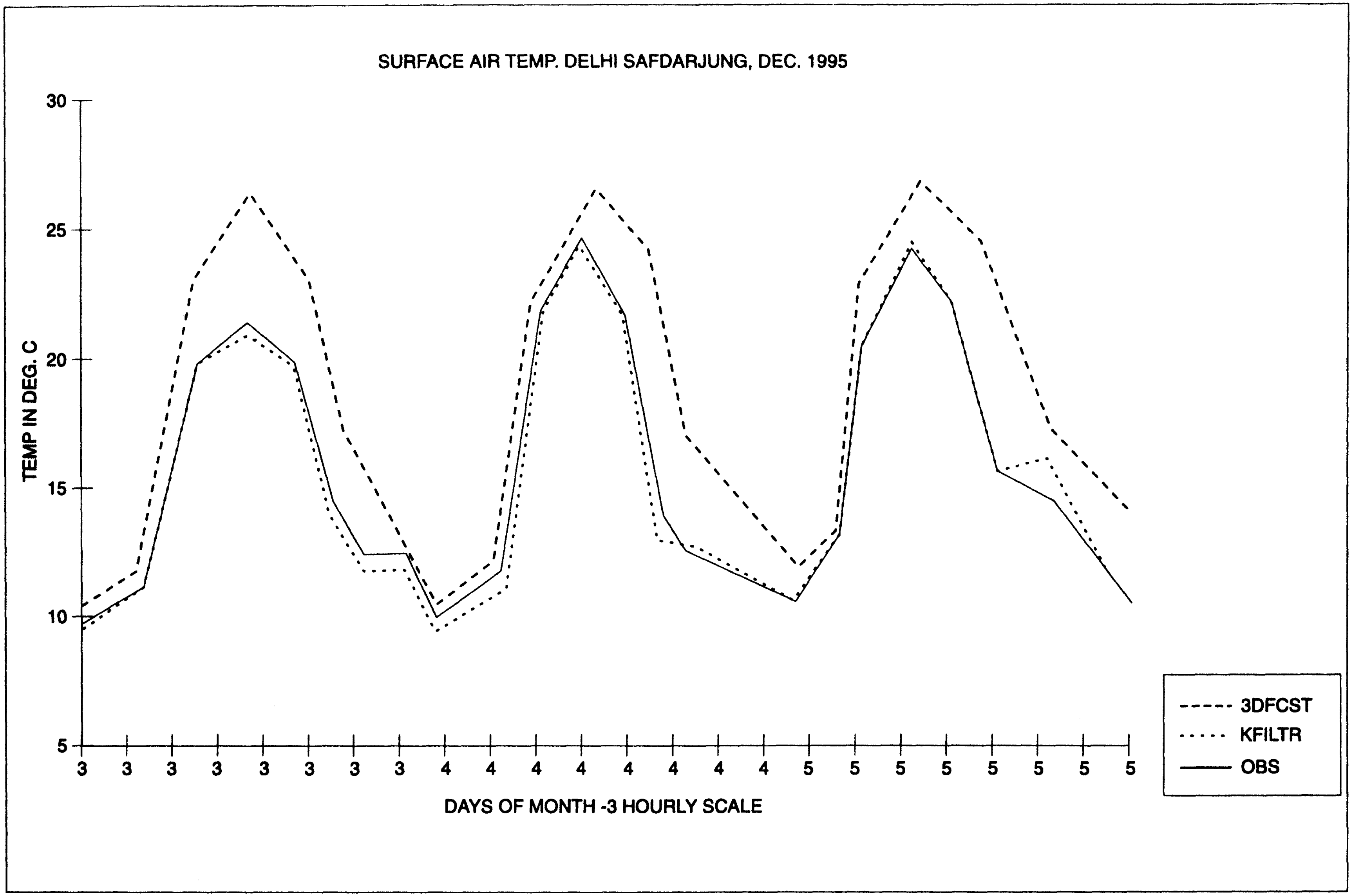

FIG. 1. Time series of corrected and uncorrected 3-day 1.36 m temperature forecasts are shown, along with the verifying observations. Solid line denote observations, dashed lines the Direct Model Output, and dash-dotted lines the forecast after correction by the Kalman filter. Results for days 3, 4 and 5 of December 1995 are shown, with 3-hourly resolution.

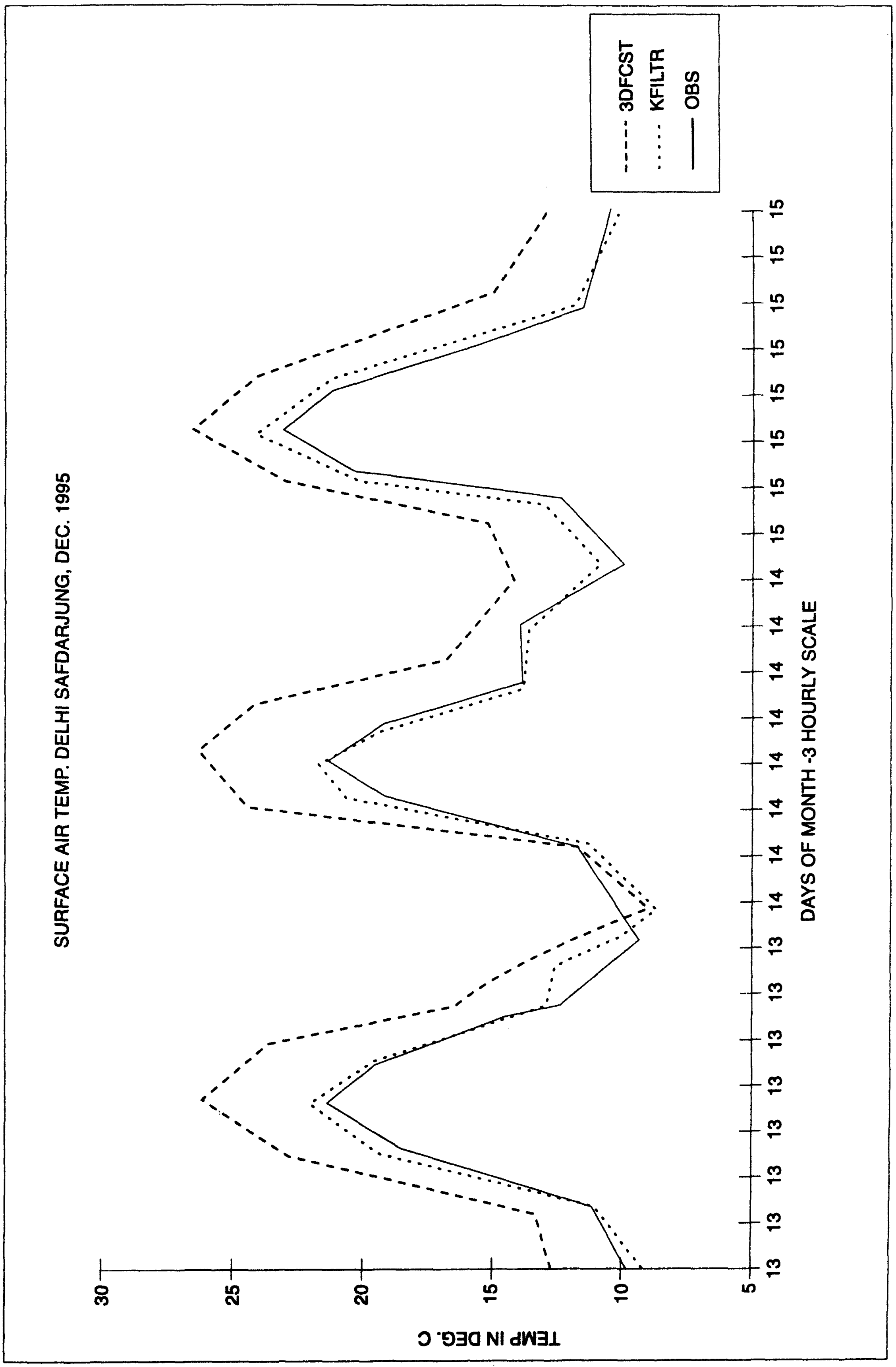

FIG. 2. Same as Figure 1, for days 13, 14 and 15 of December 1995.

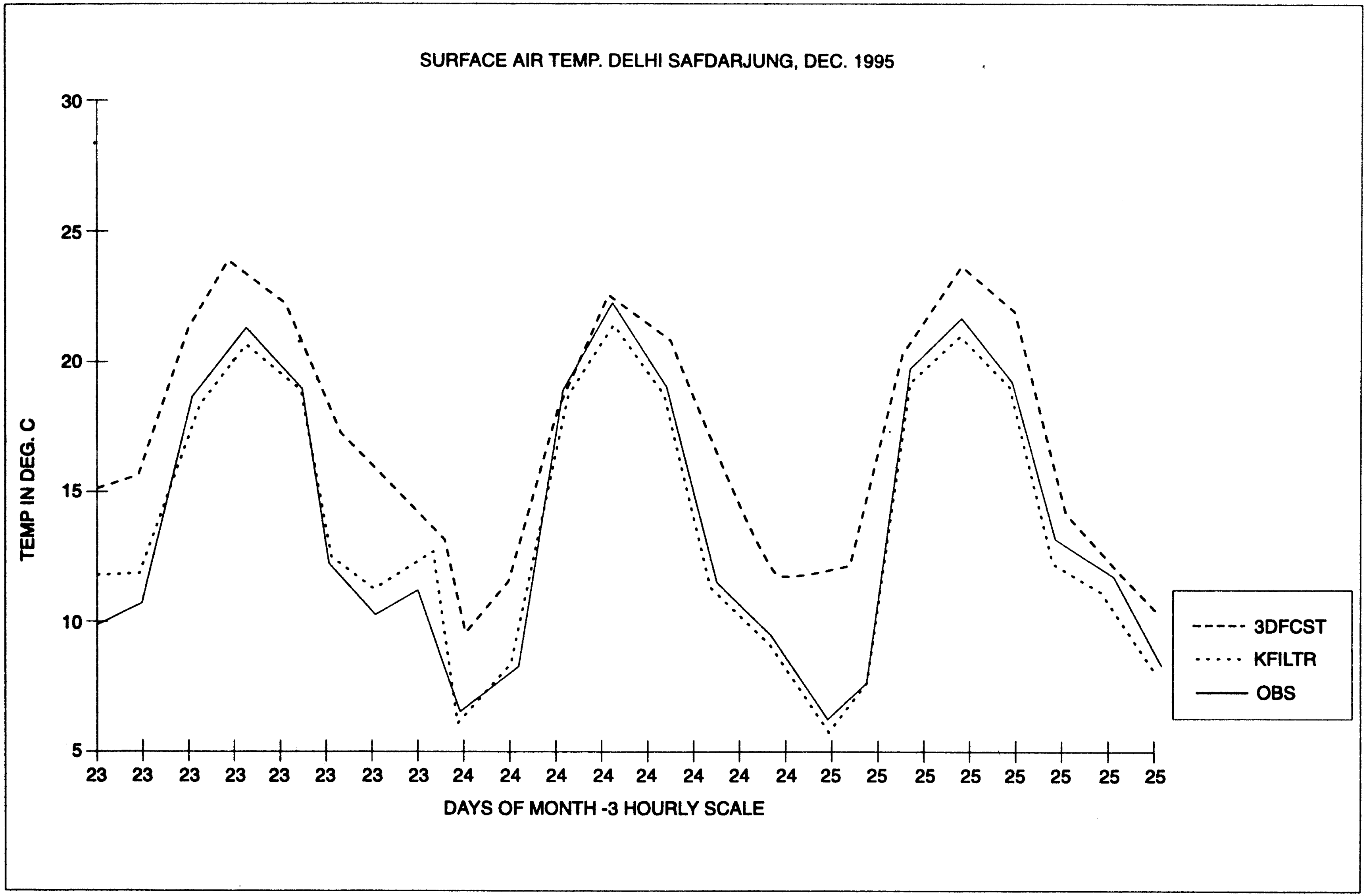

FIG. 3. Same as Figure 1, for days 23, 24 and 25 of December 1995.

few days, like day 4 (00Z) or 25 (00Z), the forecast is overcorrected while on day 5(06Z) the forecast is undercorrected. Less extreme situations are on days 24 (00Z) and 13 (09Z). It is interesting to see that overcorrections usually occur during the beginning and end of the month, while undercorrections usually occur during the middle of the month. In any case, the difference between corrected and uncorrected forecast never exceeds 2 ^{o}C. The correction is always negative if there is any; i.e., the corrected forecast temperature is never higher than the DMO forecast temperature, and this shows the basic correctness of the scheme. On some occasions, however, the corrected forecast time series is different from the original forecast time series around the maxima or minima of the diurnal cycle, because of spurious "kinks" that are present in the observed time series, which are not present in the original time series. In case of disagreement between DMO forecast and observation (due to either phase error between the two, or oversmoothing of the DMO forecast or both), the Kalman filter seems to follow the observations rather faithfully.

Figs. 4 and 5 show the maximum and minimum temperatures respectively, for each day of December 1995. As before, the original forecast, the Kalman filter corrected forecast and the observations are displayed. It is seen that the Kalman filter does not always improve the DMO forecast; although it follows the general trend of observations the agreement is not good. Here the problem is that the reported maximum and minimum temperature observations are not the same as the daily maximum and minimum temperature computed from the observation time series shown in Figs. 1-3. The observed maximum and minimum temperatures in Figs. 4 and 5 are separate observations from the maximum-minimum thermometers kept at the station at 1.36m. These are continuous observations and, therefore, can occur in reality between the valid times of forecast. We are exploring ways to solve this problem.

4. DISCUSSION

The benefits of using a Kalman filter for surface forecast correction were earlier demonstrated by using a simplified Kalman filter (Persson, 1991; Kilpinen, 1992).[11 & 12] In this work the correction was implicitly modeled as a bias (i.e., a one parameter model). Persson (1991[12]) indicated that better results may be obtained, if the correction is modeled to depend on other variables also, such as, the 850 hPa temperature and 10m wind and cloudiness in a linear manner (i.e., a two parameter model). These variables may even belong to an earlier forecast time. However, Persson's formulation of the problem was different from ours; an extension of our formulation to a multi-parameter one is not trivial. There is also the practical problem of determining just which meteorological variables affect a given station's surface temperature from a large dataset. In the opinion of Homleid (1995)[8] it might not be worthwhile to do so. Homleid (1995)[8] gave two reasons — first, the relationships may be nonlinear and, second, the Kalman filter scheme does not work well for sudden changes in weather. However, the second point was not found to be a serious problem for our case, presumably because tropical weather is generally steadier than that in mid-latitudes.

Specification of the system and observational noise matrices, particularly the relation between off-diagonal elements of the system matrix is an area which can be improved upon. In the absence of any physical guidance, it is not easy to make a stipulation. There is probably a non-zero correlation between corrections at different valid times, but it is not easy to quantify them. In fact, there is no certainty that such correlations are time independent.

The ratio of system to observed variances was determined by trial and error. Too high a value causes a fast response of the Kalman filter to rapidly changing weather, but the stability of the filter in such cases is small. A bad or unrepresentative observation can throw the system off balance. On the other hand, too low a value gives very slow response of the filter to changing weather.

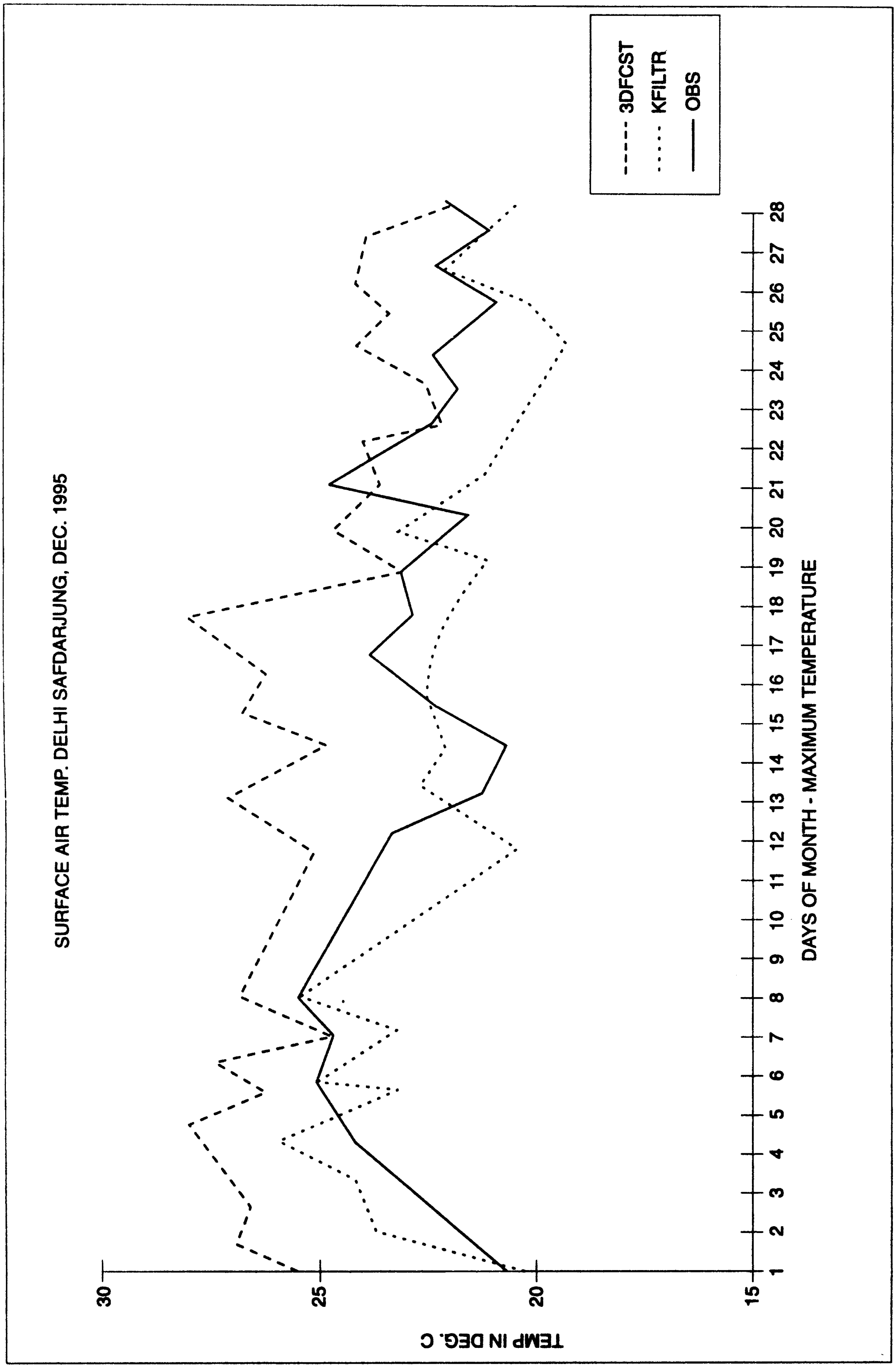

FIG. 4. Time series of daily maximum temperature at 1.36m are shown. Solid lines denote verifying observations (from maximum-minimum thermometer), dashed lines are computed from the Direct Model Output, and dash-dotted lines are computed from the forecast after correction by the Kalman filter.

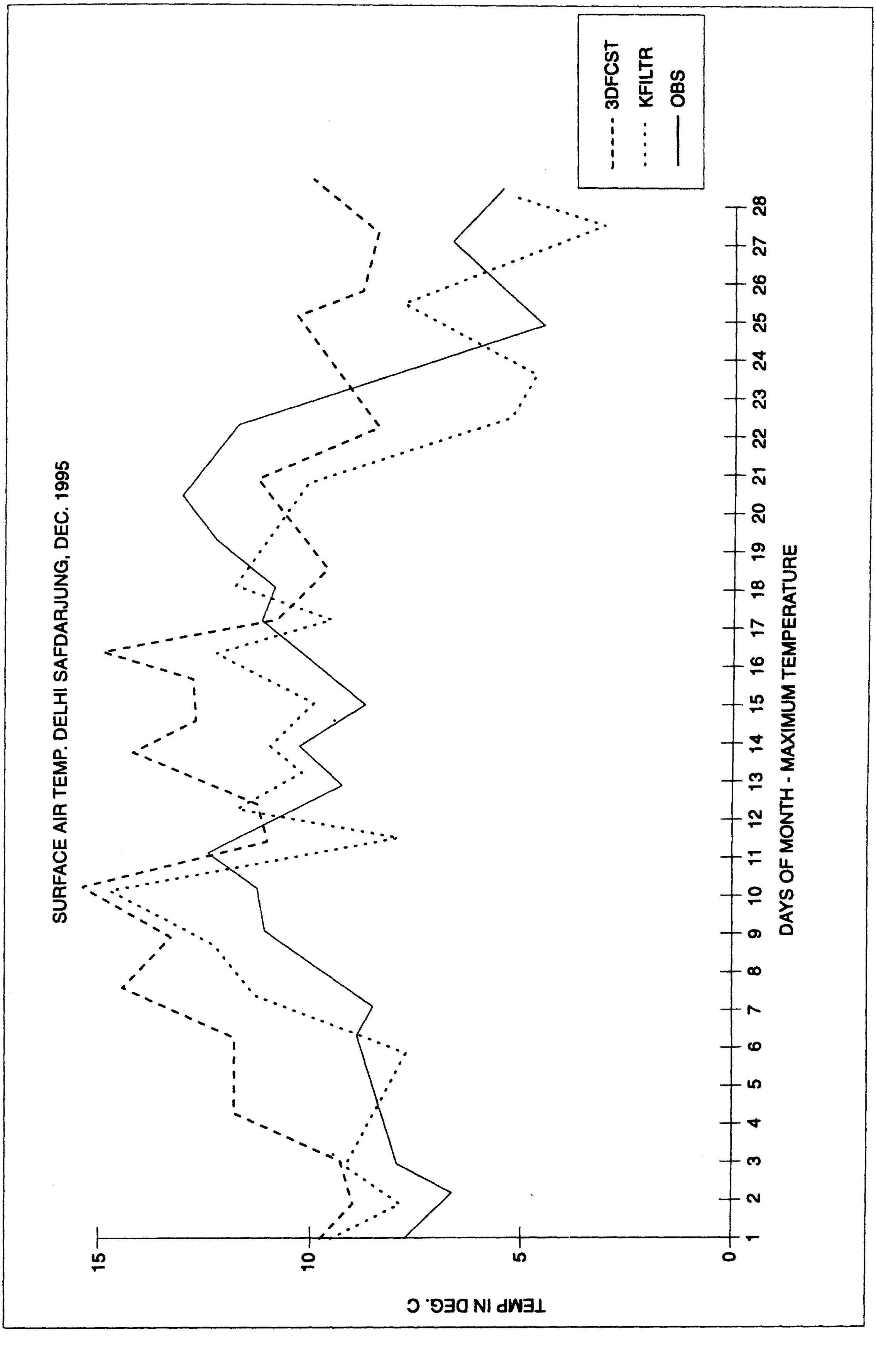

FIG. 5. Same as Fig. 4, but for the minimum temperature at 1.36m.

Apart from giving better results — from experience at NCMRWF as well as other Weather Forecasting Centres (Persson, 1991; Kilpinen, 1992)[11 & 12] — the Kalman filter scores over conventional statistical interpretation techniques in other important ways. Thus, the filter can easily adapt to seasonal and model changes, there is no need for large historical data samples and new stations can be easily integrated into the system.

An extension of the present work, apart from trying a two parameter linear model or higher parameter nonlinear models would be to perform similar exercises for other stations with different geographical conditions and also for other seasons. This is necessary to ascertain if the tuning of system and observation error covariance matrices are geography and/or weather dependent. Also, an interesting scientific question would be to see upto what forecast lead time the present procedure remains usable.

5. ACKNOWLEDGEMENT

The author is thankful to the India Meteorological Department at New Delhi for providing the synoptic observations. Thanks are also due to Mr Ashok Kumar and Ms. Parvinder Maini for discussions on the operational agrometeorological forecast dissemination at the NCMRWF. The improvement of such forecasts provided the motivation for this work.

REFERENCES

1. D. Cattani, Application d'un filtre de Kalman pour adapter les temperatures à 2 metres fournies par le modele ECMWF aux stations meteorologiques de la Suisse. *Work. Rep. SMI N* 175, 1994.
2. R. Daley, *Meteorol. Atmos. Phys.*, **60** (1996) 119-36.
3. A. Gelb, *Applied Optimal Estimation*, MIT Press, 374 pp, 1974.
4. M. Ghil, S. E. Cohn, J. Tavantzis, K. Bube and E. Issacson, In: L. Bengtsson, M. Ghil and E. Kallen (Editors), *Dynamic Meteorology : Data Assimilation Methods*. Springer-Verlag, New York, 139-224, 1981.
5. M. Ghil, *Dyn. Atmos. Ocean.*, **13** (1989) 171-218.
6. M. Ghil and P. Malanotte-Rizzoli, *Adv. Geophys.*, **33** (1991) 141-266.
7. S. Hastenrath, *Climate Dynamics of the Tropics*. Kluwer Academic Publishers, 488 pp. 1991.
8. M. Homleid, *Weather Forecast.*, **10** (1995) 689-707.
9. A. H. Jazwinski, *Stochastic Processes and Fitering Theory*. Academic Press, New York, 376 pp, 1970.
10. T. Kailath, *IEEE Trans. Automat. Contr.*, **AC-13** (1968) 646-55.
11. J. Kilpinen, *Proc. 12th Conf. on Probability and Statistics in Atmospheric Sciences*, Toronto, ON, Canada, Amer. Meteor. Soc., 11-16, 1992.
12. A. Persson, Kalman filtering - 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 Phoenomena and Their Verification, Wageningen, the Netherlands, WMO, XX_27-XX_32, 1991.
13. S. Prasad, M. Chakraborty and H. Parthasarathy. In: *Statistics and Its Applications* (Ed. B.L.S. Prakasa Rao) INSA, New Delhi, pp. 63-94, 1995.
14. N. Wiener, Nonlinear prediction and dynamics. In: *Proc. 3rd Berkeley Symposium Math. Stat. Probability*, Vol. 3, University of California Press, PP. 247-52, 1956.