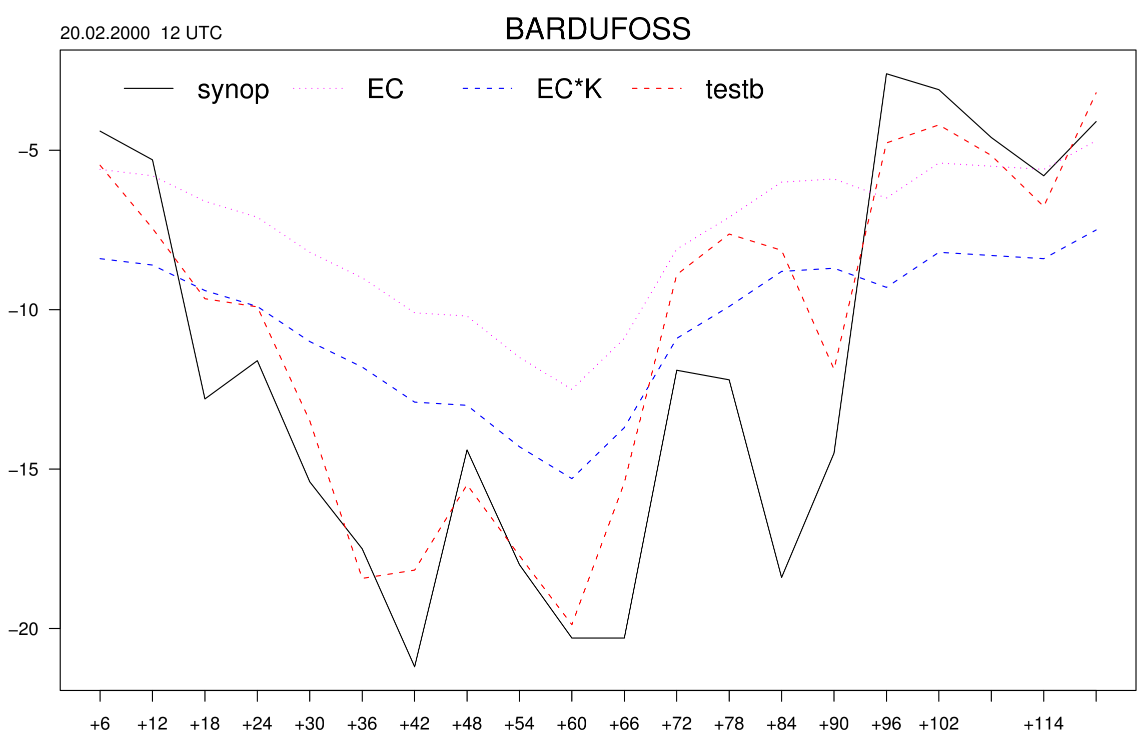
no. 6/2004
Meteorology

Weather dependent statistical adaption of 2 meter temperature forecasts using regression methods and Kalman filter

Mariken Homleid



Norwegian Meteorological Institute met.no

report

Title	Date
Weather dependent statistical adaption of 2 meter temperature forecasts using regression methods and Kalman filter	February 17, 2004
Section	**Report no.**
Meteorology	no. 6
Author	**Classification**
Mariken Homleid	● Free ○ Restricted
	ISSN 1503-8025

Abstract

Linear regression, generalized additive models, neural network and Kalman filter have been evaluated as methods to adapt numerical weather prediction forecasts to observed 2 meter temperature. The aim is to find a method and design a model which shows improved performance as compared to the operational Kalman filter procedure at the Norwegian Meteorological Institute by including weather dependencies. The temperature forecast errors are related to large scale weather parameters as wind speed and relative humidity in 850 hPa. Regression methods give good results when applied to forecasts from ECMWF. The error characteristics of these forecasts have been relatively stable the last years. The error characteristics of the Hirlam models run at the Norwegian Meteorological Institute show larger variability implying that regression methods are not well-suited. An alternative is to apply a Kalman filter to update the regressions coefficients, allowing them to adapt to changes of the systematic dependencies through the year and from one year to the next. As compared to the operational Kalman filter procedure large improvements are found in some situations. Also summary statistics show so large improvements, both for ECMWF and Hirlam10, that an updating of the operational procedure should be considered.

Keywords

2 meter temperature, statistical adaption, Kalman filter, linear regression, generalized additive models, neural networks, error characteristics of temperature forecasts

Disiplinary signature	Responsible signature
______________________	______________________

Postal address P.O Box 43 Blindern N-0313 OSLO Norway
Office Niels Henrik Abels vei 40
Telephone +47 2296 3000
Telefax +47 2296 3050
e-mail: met.inst@met.no **internet:** met.no
Bank account 7694 05 00601
Swift code DNBANOKK

1 Introduction

Statistical methods are applied in meteorology to adapt forecasts from numerical weather prediction (NWP) models to surface weather observations, e.g. temperature, wind, precipitation, cloudiness. Linear regression methods have been in use since the late sixties, (Glahn and Lowry, 1972), and are still widely in use. These methods are simple and efficient. They require a large database to estimate predictors, and work well when the error characteristics not change, or only change in a systematic manner. Increased computer resources have made it possible to apply nonlinear methods as e.g. generalized additive models (Vislocky and Fritch, 1995) and neural networks (Marzban, 2003).

Kalman filter theory (Kalman, 1960; Kalman and Bucy, 1961; Gelb, 1974; Priestly, 1981) provides an excellent tool for combining observations and forecasts to correct for systematic errors. Examples of applications are found in (Persson, 1991; Simonsen, 1991; Kilpinen, 1992). The Kalman filter approaches do not require much data back in time for model estimation. It is however useful with some data to study the performance. This means also that the Kalman filter models are robust with respect to model changes implying changes in the forecast error characteristics.

At the Norwegian Meteorological Institute a Kalman filter procedure has been applied operationally since 1992 to correct 2 meter temperature (T_{2m}) forecasts, (Homleid, 1995). The procedure is relatively simple and robust. The same procedure is applied to forecasts from different NWP models, at all observing stations. It works well in stable weather situation, but the behavior might be bad at sudden weather changes. We would like to improve the performance by including weather dependencies. Section 3 presents the operational Kalman filter procedure.

Section 4 describes experiments with regression techniques on T_{2m} forecasts from ECMWF. We find improvements as compared to the operational Kalman filter. The performance is, however, not very nice when the error characteristics change from one year to the next. The error characteristics of the T_{2m} forecasts from Hirlam vary even more, as descriebed in (Homleid and Ødegaard, 2000) and summarized in section 2. In section 5 a Kalman filter is designed to update regression coefficients. The experiences from the experiments with simple regression techniques are useful when designing the Kalman filter model.

Section 6 discuss the experiences: Are the improvements significant as compared to the operational Kalman filter?

2 Error characteristics of 2 meter temperature forecasts

When comparing temperature forecasts and observations, both systematic and unsystematic differences are found. The systematic errors can be related to shortcomings of the NWP model:

- The difference between the topographical height in the NWP model and the actual terrain height might be several hundred meters, implying large and systematic errors in the T_{2m} forecasts.
- At many coastal stations the land/sea-mask of the NWP model might introduce systematic errors. The T_{2m} forecasts are either too close to the sea surface temperature (SST) or too close to "inland

temperatures".

- At inland stations the NWP models are not cold enough when it is very cold. The soil temperatures of the NWP model are given by climatological mean values restricting the NWP temperatures to reach the lowest temperatures observed.
- The temperature is influenced by the cloud cover, and systematic errors in the cloud cover forecasts might introduce systematic errors in the T_{2m} forecasts.
- Surface properties of the NWP model can also contribute to systematic errors in the T_{2m} forecasts. The temperature in spring time is very sensitive to the snow cover.

The magnitude of the contributions to the forecast error from the error sources above depends on the weather situation and will vary from station to station and through the year.

Situations with erroneous pressure fields in the NWP model might give temperature forecasts with large errors of an unsystematic character. Other shortcomings of the NWP model might also contribute to such errors.

Generally the magnitude of the forecast errors reflects the temperature variation at a station. The errors are small at the coastal stations, with standard deviations of the errors (SDE) less than 2 °C through the year. At inland stations the monthly SDE's range from 2 °C in the summer months to 3 - 8 °C in the winter months. The mean error (ME) of a station often reflects the topographical height error of the station. The monthly mean errors often show a yearly cycle. Coastal stations with temperature forecasts too close to SST will typically have negative mean errors in spring and summer, positive mean errors in autumn and winter, see Fig. 1 from Færder light house in the Oslo fjord and Fig. 2 from Bodø. Inland stations often have positive mean errors in winter, see Fig. 3 from Bardufoss, an inland station in the northern part of Norway.

What is said so far about the forecast quality refers to both the global ECMWF model and the Hirlam models running operationally at the Norwegian Meteorological Institute, Hirlam10, Hirlam20 (from March 2002) and Hirlam50 (untill March 2002) with 10 km, 20 km and 50 km resolution respectively. Here follow some model specific comments. The quality of the ECMWF forecasts has been very stable the last years. The error characteristics of the Hirlam models show larger variations, see Figs.1, 2, 3. The differences are largest in the coastal zone, where the Hirlam temperatures are too close to the sea surface temperature and therefore very sensitive to the quality of the sea surface temperature, which is given as a parameter field. An other example of systematic errors is errors related to the cloud cover forecasts. The Hirlam10/50 cloud cover forecasts at the Norwegian Meteorological Institute were untill March 2002 based on the Sundqvist scheme (Sundquist et al., 1989) which underestimates the cloud cover, a problem which increases with increasing resolution. In March 2002 the operational Hirlam models were updated to Hirlam 5.3.1, with the STRACO scheme for cloud parametrization (Sass, 1997). This scheme overestimates the cloud cover at many locations. The changes in cloud cover characteristics will imply changes in the error characteristics of the temperatures.

3 The operational Kalman filter procedure

2 meter temperature forecasts have since 1992 been operationally corrected by a Kalman filter procedure described in (Homleid, 1995). Here follow a short presentation of the notation, model design and some comments on the performance. General Kalman filter theory might be found in e.g. (Gelb, 1974) or (Priestly, 1981).

Let $\mathbf{x}_t$ be a vector describing the state of a process at time t. The state at time t is related to the state at time $t-1$ through the

system equation:

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

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

The state $\mathbf{x}_t$ is not observable but is related to the observations $\mathbf{y}_t$ through the

observation equation:

$$\mathbf{y}_t \ = \ \mathbf{H}_t\mathbf{x}_t \ + \ \mathbf{v}_t \ , \quad t = 1, ..., T \tag{2}$$

where $\mathbf{H}_t$ is the observation matrix and $\mathbf{v}_t$ is a vector with the random observation error.
$\mathbf{w}_t$ and $\mathbf{v}_t$ are Gaussian zero mean white noise processes with covariance matrixes $\mathbf{W}_t$ and $\mathbf{V}_t$.

The Kalman filter theory gives an algorithm for recursively estimating $\mathbf{x}_t$, t=1,...,T utilizing all observations up to time t. In our case $\mathbf{x}_t$ is a vector of length 8 with the systematic deviations between temperature forecasts and observations at the observing times 00, 03, 06, 12, 15, 18 and 21 UTC. The estimate of $\mathbf{x}_t$ is used to correct forecasts produced at time t. The Kalman filter procedure thus produces 8 diurnal varying corrections which will be weighted means of the last days differences between T_{2m} forecasts and observations, with weights decaying back in time. The rate of decay depends on the relation between the variances W_t and V_t. The correlation in time between the 8 diurnal corrections follows from the specification of the covariance matrix $\mathbf{W}_t$.

The specified procedure works well in stable weather situation. The summary statistical results show monthly mean errors close to zero. The behavior at sudden weather changes is, however, not so good. The largest errors occur at inland stations in winter. Long periods with low temperatures and forecasts that are not cold enough will give large positive corrections. When the temperature increases again, the corrected temperature forecasts will often be much too high. These errors are reduced by an 'ad hoc' procedure which set the corrections to values close to the monthly mean errors when the temperature forecasts increase more than 5 °C in 24 hours. A more general handling of weather dependencies would probably give better results. It is, however, not straightforward to extend the operational Kalman filter procedure to take weather dependencies into account. In section 5 we follow an approach closer to the ones of (Persson, 1991; Simonsen, 1991; Kilpinen, 1992).

4 Regression methods

Linear regression models have been fitted to adapt NWP forecasts to weather observations since the late sixties. The later years increased computer resources have made it possible to apply nonlinear methods to a large extent. This section discuss experiments performed with linear regression models, generalized additive models and neural networks to adapt ECMWF forecasts to observed 2 meter temperatures.

4.1 Linear regression

A linear model can be formulated by

$$Y = \alpha + \sum_{j=1}^{p} \beta_j X_j + \epsilon \tag{3}$$

The observed temperature, Y, is supposed to be a linear combination of a selection of NWP forecasted parameters, X_j, the predictors. The forecasted 2 meter temperature is often one of the predictors. When the predictors are chosen, the regression coefficients, α and β_j, are estimated by least-squares from data sets with observations and forecasts, training data. ϵ is supposed to be Gaussian zero mean white noise with constant variance. The model is evaluated by making predictions on data sets not included in the training set and calculate summary statistics on the differences between the predicted and the observed temperature. The challenge is to find an optimal set of predictors. Experiences show that too many predictors may lead to over fitting. Simple models with only a few predictors are more robust.

The temperature forecast errors depend on the weather situation, as discussed in section 2. Which NWP parameters and which levels of the atmosphere are good indicators of the large scale weather? In traditional MOS technique, the predictors might be chosen automatically by a screening regression procedure (Glahn and Lowry, 1972). Physical considerations might also be important for the choice of parameters and levels. In Norway the largest forecast errors occur at inland stations in winter time, in situations with clear sky and no wind. Cloud amount is deduced from relative humidity in 925, 850 and 700 hPa. As the forecasts of relative humidity is superior to the forecasts of cloud amount, the relative humidity is preferred as a predictor. The relative humidity in 850 hPa seems to be the best indicator of the presence of clouds and the amount of radiation emitted upward, conditions which are of importance for the 2 meter temperature.

A linear model has been fitted with the S-Plus function lm (Venables and Ripley, 1994). Predictors under consideration have been forecasts of 2 meter temperature, temperature in 700 hPa, relative humidity and wind speed in 850 hPa, surface pressure and total surface radiation.

When fitting a model for all observing times simultaneously, the forecast valid time has to be taken into account in some way, as the forecast errors often show systematic diurnal variation. A cosine term, cos((TIM-1.5)*pi/12), is included to handle a possible diurnal variation in the forecast errors. To capture the variation through the year, monthly sets of predictors are estimated. Predictors to be used in e.g. May are estimated from data from April to June the previous year. Predictions of +6, +12, +18, +24 forecasts are based on predictors estimated from 6, +12, +18, +24 forecasts, predictions of +30, +36, +42, +48 forecasts are based on predictors estimated from +30, +36, +42, +48 forecasts,.....

Several experiments have been performed to explore the possibilities with linear regression models. The results of the experimental procedures have been compared to the direct output NWP forecasts (ec) and the operationally Kalman filter corrected forecasts ec*K. The above model performs well on Kautokeino data in winter time, giving lower temperatures in cold time periods and nicer performance than ec*K when the weather changes. The improvements are reflected in the summary statistics showing significantly reduced SDEs. Fig. 4 shows results from an experiment where the estimation of predictors are based on 1999 data, the training data set, and predictions are made for 1999 and 2000. Not surprisingly the best performance is obtained in 1999, with ME close to zero. In 2000 the monthly ME's are larger, especially in March and November. In summer time when the error are much smaller no significant improvements are seen.

Corresponding results are presented for Bardufoss 1999-2000 in Fig. 5. The predictors are estimated from 1999 data. Again we find larger ME and lower SDE than obtained with the operational Kalman filter.

The results presented so far are for predictions with predictores estimated from 3 months data sets. Is it possible to improve the results by utilizing more data when estimating the predictors? Predictions for the years 1997 to 2000 with predictors based on data from these years except the prediction year have been performed for some stations.
Three linear models with different sets of predictors are compared:
pred1 - ec.T_{2m}, ec.TT.700 and cos((TIM-1.5)*pi/12)*ec.RH.850)
pred2 - ec.T_{2m}, (ec.TT.1000-ec.TT.700) and cos((TIM-1.5)*pi/12)*ec.RH.850)
pred3 - ec.T_{2m}, obs.T_{2m}, (ec.TT.1000-ec.TT.700) and cos((TIM-1.5)*pi/12)*ec.RH.850)

pred2 is as pred1 but uses the vertical temperature gradient dT=ec.TT.1000-ec.TT.700 instead of ec.TT.700
pred3 is as pred2 but includes also the latest temperature observation as a predictor

Fig. 6 gives results for Bardufoss for 1999, with predictors based on 1997,1998,2000 data
Fig. 7 gives results for Bardufoss for 2000, with predictors based on 1997,1998,1999 data

Corresponding results have also been produced for Tromsø and Bodø. We find that predictions based on larger data sets generally give better results. The inclucion of the latest observation (pred3) improves the results slightly, whereas the results of pred1 and pred2 are very close.

The results presented so far are monthly mean values for 12+6, 12+12, 12+18 and 12+24 forecasts. It is also interesting to study the results as a function of forecast length. Fig. 8 gives summary statistics for Bardufoss for January and February 2000 for 12+6, +12, +120 forecasts, and Fig. 9 gives corresponding results for spring 2000. The diurnal oscillations in the ME's reflects that the observed temperatures have larger diurnal amplitudes than the forecasts. Both the Kalman filter and the linear models increase the diurnal amplitudes and have ME's close to zero.

There is large variation in the results from station to station, through the year and from one year to the next, but generally
- ec*K has lower ME's than pred
- pred has lower SDE's than ec*K

- there are larger improvements with both ec*K and pred at locations and in situations with large errors (inland, winter)

4.2 Generalized additive models

The dependence of the temperature forecast errors on predictors as relative humidity and wind speed is not linear. Generalized additive models (Hastie and Tibshirani, 1990) have been designed to handle non-linear dependencies. We will see if it is possible to improve the results obtained with linear models.

A generalized additive model can be formulated by

$$Y = \alpha + \sum_{j=1}^{p} f_j(X_j) + \epsilon \tag{4}$$

where f_j are smooth functions and the definition of the other variables follows (3).

The S-Plus function gam (Venables and Ripley, 1994) is applied to fit a generalized additive model with predictors:
ec.T_{2m}, ec.TT700, s(ec.RH850), s(ec.FF850), s(I(ec.P-1000)) and cos((TIM-1.5)*pi/12)
s specifies a smoothing spline.

The model is fit using a local scoring algorithm, which iteratively fits weighted additive models by backfitting. The backfitting algorithm is a Gauss-Seidel method for fitting additive models, by iteratively smoothing partial residuals. The algorithm separates the parametric from the nonparametric part of the fit, and fits the parametric part using weighted linear least squares within the back-fitting algorithm.

We find larger ability to predict extreeme temperatures than with linear modes but also more examples of overshooting. The results are discussed in section 4.4.

4.3 Neural networks

Neural networks represent another generalization of linear regression functions. S-Plus provides software to fit feed-forward neural networks with one hidden layer, nnet, (Venables and Ripley, 1994). This function has been applied to Kautokeino data. Forecasts of T_{2m}, temperature in 700 hPa, relative humidity and wind speed in 850 hPa and a cosine transformation of the forecast time were chosen as predictors.

It seems possible to obtain better results with neural networks than with linear models. But we encountered the problem that repeated fits of the model gave predictions which might disagree with several degrees. This problem might be handled by taking the mean of repeated predictions.

4.4 Discussion of results obtained with regression methods

Generalized additive models and neural networks show larger capability than linear models of predicting extreeme temperatures. Unfortunately, the predicted temperatures might be too extreeme in some situations. Large temperature forecast errors of an unsystematic character will disturb the parameter estimation and lead to such problems. The better and more stable the quality of the NWP model is, and the more data, the larger are the possibilities to establish the relationship between predictors and observed temperatures.

(Marzban, 2003) demonstrates promising results obtained when postprocessing 2 meter temperature forecasts with a neural network, and gives lot of useful information for an implementation of the method. He indicates that overfitting which is often a large problem for nonlinear models is less of a concern for neural networks where the number of free parameters are comparable with linear models even if the possibilities to handle non linearities are larger.

5 Linear regression with Kalman filter updated coefficients

Regression methods work well when the quality of the forecasts are relatively stable, e.g. on the ECMWF model output. Our aim is, however, to develop an opertional procedure for postprocessing forecasts from several NWP models, including models with larger variation in the forecast error characteristics. The regession coefficients might alternatively be updated by a Kalman filter (Persson, 1991; Simonsen, 1991; Kilpinen, 1992). The regression coefficients will then be adapted to the forecasts and observations of the last days/weeks. It is, however, useful with data for some years to design and evaluate the Kalman filter model. Section 5.1 discuss a Kalman filter model applied to ECMWF forecasts to see if it is possible to get results comparable with those obtained with regression techniques. An identical Kalman filter model has also been evaluated for Hirlam10 forecasts. The results are discussed in section 5.2.

5.1 ECWMF forecasts

When designing a Kalman filter model to update regression coefficients, the degrees of freedom are unlimited. Firstly, a set of predictors has to be chosen. Then the error covariance matrixes have to be specified, either by a statistical estimation procedure or by ’tuning’ them to make the Kalman filter work as wanted.

Experiments with various predictors have been performed. The ’best’ results were obtained with forecasts of 2 meter temperature, relative humidity and wind speed in 850 hPa - ’best’ in the sense that it was not possible to improve the summary results significantly with other sets of predictors. The inclucion of RH_{850hPa} and FF_{850hPa} increases the ability to predict the very cold temperatures at inland stations in winter time, but might lead to large errors when these forecasts are wrong. A transformation of the forecasts improves the performance. RH_{850hPa} has been standardized by yearly values of mean and standard deviations, T_{2m} by running means of monthly values. A logarithmic transformation of FF_{850hPa} before standardization increases the sensibility to the low wind speeds.

Here follows a specification of a model with T_{2m}, RH_{850hPa} and FF_{850hPa} as predictors. The notation is the same as in section 3.

The regression coefficients make up the state vector: $\mathbf{x}_t = [\ \alpha_t\ \ \beta_{1t}\ \ \beta_{2t}\ \ \beta_{3t}\ ]$
The system matrix: $\mathbf{F}_t = \mathbf{I}$
The observation vector: $\mathbf{H}_t = [\ 1\ \ f_T(T_{2m})\ \ f_U(RH_{850hPa})\ \ f_F(FF_{850hPa})\ ]$

T_{2m}, RH_{850hPa} and FF_{850hPa} are 00+6,+12, 12+6,+12 forecasts valid at time t. f_T, f_U and f_F indicate that the forecasts have been transformed, as explained above. There is one Kalman filter for each observing time, as the dependence of the temperature forecast errors on the predictors might vary through the day. Predictors valid at e.g. 12 UTC are used to predict the temperatures valid at 12 UTC on the following days for the respective forecast lengths. This approach might give regression coefficients which hopefully show a physical variation through the day. Unfortunately, there might be a random variation of the regression coefficients through the day leading to unphysical features of the corrected temperature curves. Such problems occur more often when the regression coefficients are allowed to change much in one time step. The rate of change is controlled through the specification of the error covariances. Several candidates have been tested, among them:

testa: $W_{1,1} = 0.1$, $W_{i,i} = 0.1 * 0.04$, i = 2,3,4 , $W_{i,j} = 0$, $i \neq j$
testb: $W_{1,1} = 0.01$, $W_{i,i} = 0.01 * 0.04$, i = 2,3,4 , $W_{i,j} = 0$, $i \neq j$
testc: $W_{1,1} = 0.001$, $W_{i,i} = 0.001 * 0.04$, i = 2,3,4 , $W_{i,j} = 0$, $i \neq j$
testd: $W_{1,1} = 0.0001$, $W_{i,i} = 0.0001 * 0.04$, i = 2,3,4 , $W_{i,j} = 0$, $i \neq j$

These covariance matrixes are all designed to allow the first regression coefficient to change more rapid in time than the coefficients of T_{2m}, RH_{850hPa} and FF_{850hPa}, as they are supposed to vary slowly through the year. Figures 10 and 11 are included to illustrate the performance of testb at Bardufoss on the 20th of February 2000 and the 26th of February 2001. Experimentally corrected 12+6, +12, ... , +120 forecatsts (testb) are compared to observations, uncorrected ECMWF forecasts (EC) and operationally corrected forecasts (EC*K) at the top. Transformed forecasts of T_{2m}, RH_{850hPa} and FF_{850hPa} are given at the bottom, and the contribution from the different predictors to the final correction in the middle. It can be seen that low wind speeds and also low values of relative humidity contribute to reduced temperatures. In Fig. 10 only dT1, the contribution from the constant term, α_t, shows a diurnal variation. The ECMWF model seems to give a good forecast of the weather development for the following days and the result of the experimental Kalman filter procedure is very nice. In Fig. 11 also the term related to the relative humidity, dT3, shows diurnal variation, giving lower temperatures in night time when the relative humidity is low. testb shows improved results on the first four days, when compared to EC and EC*K, but on the fifth day the temperature increases much more than forecasted and testb is much too low. Such larger errors occur relatively frequent when the weather development of the NWP model is incorrect.

Summary statistics as a function of forecasts length for different time periods have been studied for a selection of stations to compare the results of experimentally and operationally Kalman filter corrected and uncorrectes ECMWF forecasts. As an example results for Bardufoss January and February 2000 are shown in Fig. 12. In this time period the ECMWF forecasts are too smooth, with standard deviations below 4 °C while the standard deviations of the observations vary between 6 and 7 °C. The standard deviations of the forecasts are too low mainly because the temperatures are not low enough in very cold situations. The standard deviations of testa are close to the observed , the ones of testc are comparable with the ones of the operationally Kalman filter corrected and the ones of testb are between the values of testa and testc. While testa and testb reduce the mean errors to values close to zero, testc gives to slow response and introduce negative mean errors. The standard deviations of the errors show significant improvements with testa and testb as compared to the uncorrected and operationally Kalman filter corrected forecasts for +6, ... , +96 forecasts. These results are typical for inland stations in winter time. For Bardufoss EC*K is often as good as the experimental versions the rest of the year. Summary statistics for 2000 and 2001 show only minor improvements of the experimental versions for Bardufoss, see Fig. 13. The results vary through the year and from station to station. There are only minor changes in the summary statistics for many stations. The largest improvements are found at stations in the south-eastern part of Norway, at the light house Færder, and two stations in Oslo, Blindern and Tryvasshøgda, see Figs. 14, 15, 16. The summary results of testb are slightly better than the results of testa and testc. The results obtained with testd are far too smooth.

5.2 Hirlam10 forecasts

Experiments with Kalman filter updated regression coefficients, parallell to the experiments desriebed in section 5.1, have also been performed with Hirlam10 temperature forecasts. Hirlam10 data is availables as fields on disk from 1999, making it possible to evaluate the performance on all Norwegian synoptic stations. The results at 37 representative stations have been studied carefully. Different sets of predictors and covariances have been evaluated. Again T_{2m}, RH_{850hPa} and FF_{850hPa} were finally chosen as predictors and again the covariances defined by testb in section 5.1, seem to give the best results. The largest improvements as compared to the operational Kalman filter corrected forecasts were found at stations in the fjords, e.g. Vadsø and Tafjord. Minor improvements are found at many stations, see e.g. results for Bardufoss, Tromsø, Bergen and Oslo, Figs. 17, 18, 19, 20. At stations at the coast the operational Kalman filter is often slightly better than the experimental versions. Summary results for 37 Norwegian stations show significant improvements of the experimental versions, as compared to operational Kalman filter, Fig. 21.

6 Conclusions and discussion

Experiments have been performed with linear regression models, generalized additive models and neural networks to adapt ECMWF forecasts to observed 2 meter temperatures. Generalized additive models and neural networks show larger capability than linear models to predict extreme temperatures. This might give nice perfomance when the weather develpoment of the NWP model is good, but large errors when not. The better and more stable the quality of the NWP model is, and the more data, the larger are the possibilities to establish good relationships between predictors and observed temperatures. It should be possible to improve the results obtained with neural networks on ECMWF data. The potential of the method has not been fully explored in this report.

Experiments with Kalman filter updated regression coefficients have been performed on both ECMWF and Hirlam10 forecasts. The regression coefficients adapt when the error characteristics varies, and good results are found both for Hirlam10 and ECMWF forecasts. The results for the ECMWF forecasts are comparable with the results obtained with linear regression.

The results obtained with Kalman filter updated regression coefficients are so good as compared to the results of the operational Kalman filter that an operational implementation should be concidered. This report focus on the evaluation by summary statistical measures. As the corrected temperature forecasts are often presented graphically as 'meteograms' it is also important that the corrections show a nice and smooth performance. Some way of smoothing of the final corrections is needed in an operational implementation.

The quality of the T_{2m} forecasts from ECMWF is superior to the quality of Hirlam forecasts and the error characteristics are stable. It could be interresting to go one step further and produce probability forecasts, following an approach with local quantile regression as desciebed in (Bremnes, 2004). He applies the method on precipitation.

An additional way of improving the quality of T_{2m} forecasts from Hirlam should be mentioned, although the potential improvements are not investigated at this stage. The T_{2m} forecasts from Hirlam are averages of surface dependent temperatures calculated for different surface types in each grid point. The surface types are bare ground, low vegetation, forrest, sea and ice. It should be possible to optimize the weighting and e.g. reduce the systematic errors in the coastal sone. This would lead to improved results not only in the observation points.

References

Bremnes, J. B., 2004: Probabilistic forecasts of precipitation in terms of quantiles using NWP model output. *Monthly Weather Review*, **132**, 338–347.

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

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.

Hastie, T. and R. Tibshirani, 1990: *Generalized Additive Models*. Chapman and Hall, London.

Homleid, M., 1995: Diurnal corrections of short-term surface temperature forecasts using the Kalman filter. *Weather and Forecasting*, **10**, 689–707.

Homleid, M. and V. Ødegaard, 2000: A study of 2 meter temperature forecasts from NWP models at Norwegian synop stations. Technical Report 112, Norwegian Meteorological Institute.

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

Kalman, R. E. 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*, Amer. Meteor. Soc., Toronto, ON, Canada, 11–16.

Marzban, C., 2003: Neural networks for post-processing model output: ARPS. *Monthly Weather Review*, **131**, 1103–1111.

Persson, A., 1991: Kalmanfiltering - a new approach to adaptive statistical interpretation of numerical meteorlogical 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*, WMO, Wageningen, the Netherlands, XX–27 – XX–32.

Priestly, M. B., 1981: *Spectral analysis and time series*. Academic Press.

Sass, B. H., 1997: Reduction of numerical noise connected to the parameterization of cloud and condensation processes in the HIRLAM model. HIRLAM Newsletter 29.

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*, WMO, Wageningen, the Netherlands, XX–33 – XX–37.

Sundquist, H., E. Berge, and J. E. Kristjansson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Monthly Weather Review*, **117**, 1641–1657.

Venables, W. N. and B. D. Ripley, 1994: *Modern Applied Statistics with S-Plus*. Springer-Verlag.

Vislocky, R. L. and J. M. Fritch, 1995: Generalized additive models versus linear regression in generating probabilistic MOS forecasts of aviation weather parameters. *Weather and Forecasting*, **10**, 669–680.

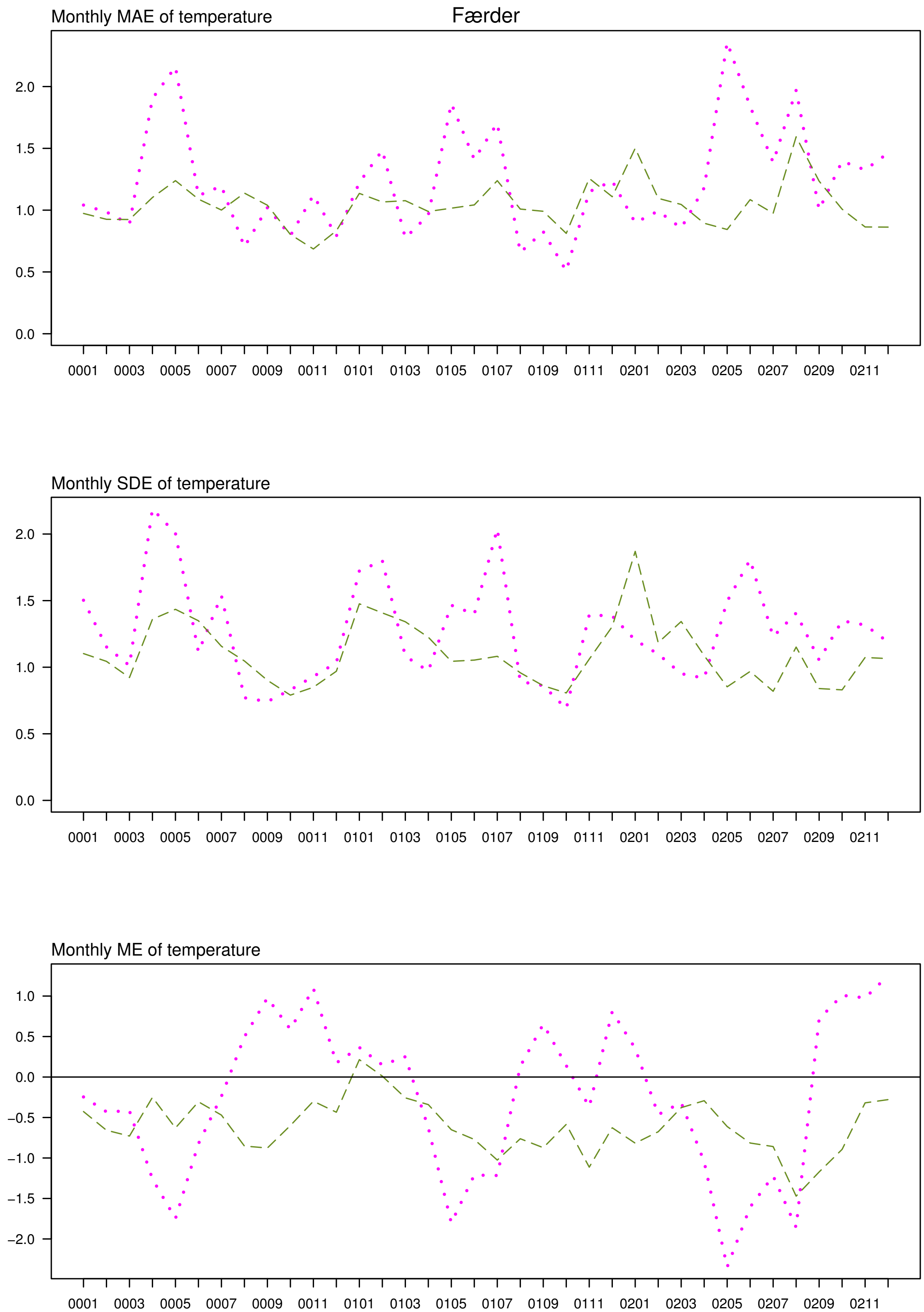


Figure 1: Færder 2000-2002; monthly MAE, SDE and ME of 12+6,+12,+18,+24 temperature forecasts from Hirlam10 (pink dotted) and ECMWF (grey dashed).

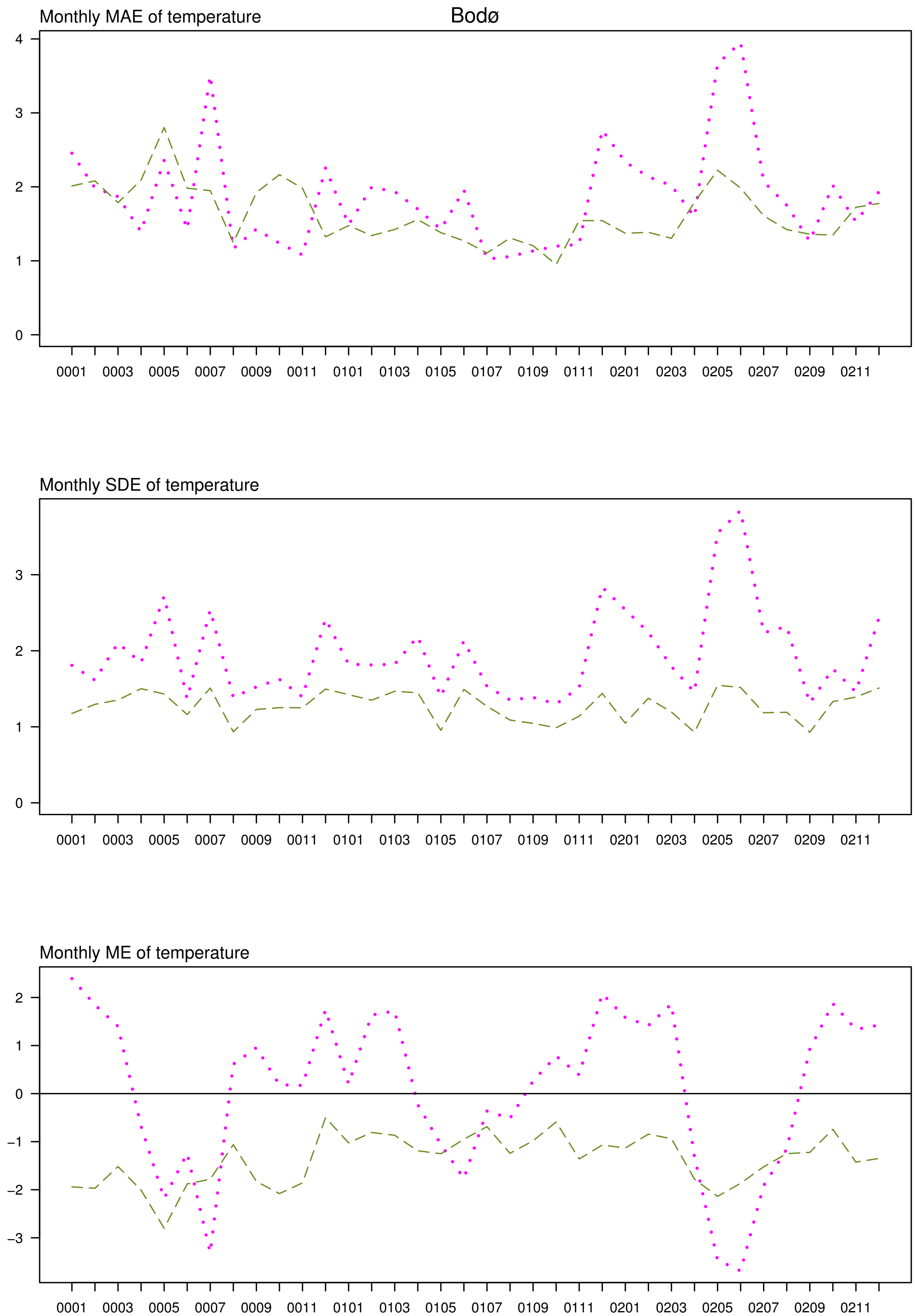


Figure 2: Bodø 2000-2002; monthly MAE, SDE and ME of 12+6,+12,+18,+24 temperature forecasts from Hirlam10 (pink dotted) and ECMWF (grey dashed).

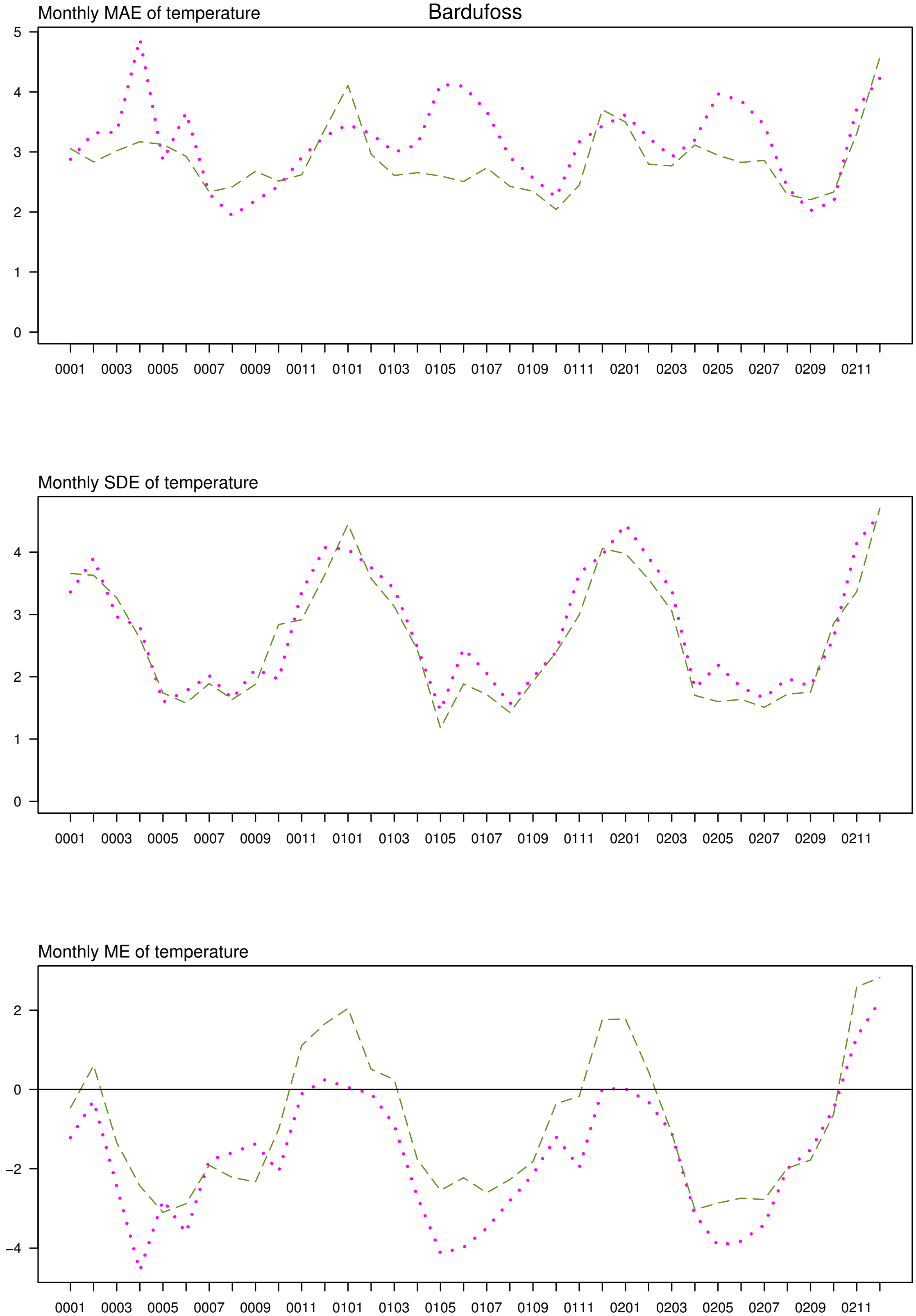


Figure 3: Bardufoss 2000-2002; monthly MAE, SDE and ME of 12+6,+12,+18,+24 temperature forecasts from Hirlam10 (pink dotted) and ECMWF (grey dashed).

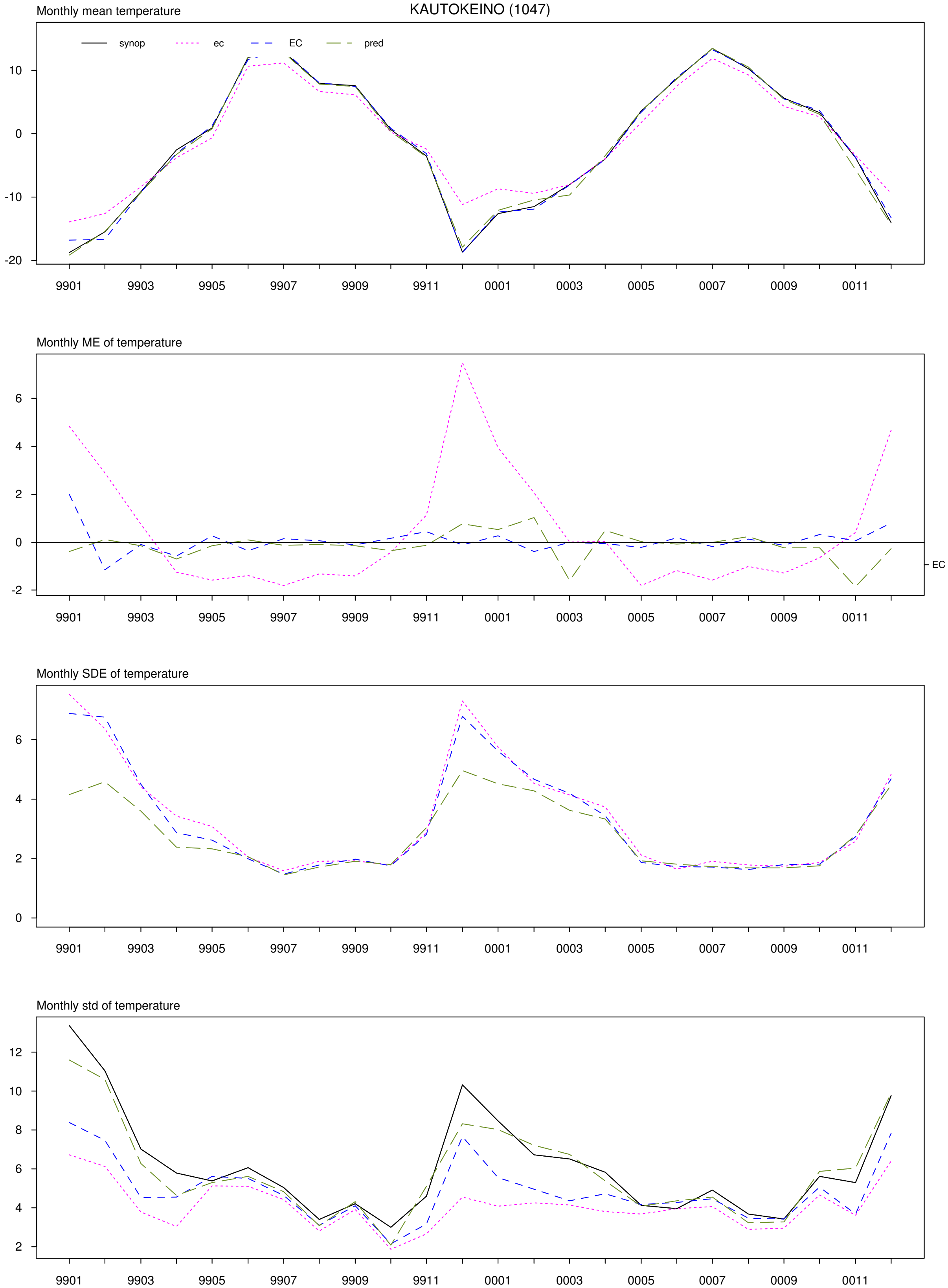


Figure 4: Kautokeino 1999-2000; monthly mean, ME, SDE and standard deviation of observations (black lines) and 12+6,+12,+18,+24 temperature forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed), predictions with linear regression based on 1999 data (grey long dashed).

Figure 5: Bardufoss 1999-2000; monthly mean, ME, SDE and standard deviation of observations (black lines) and 12+6,+12,+18,+24 temperature forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed), predictions with linear regression based on 1999 data (grey long dashed, red long dashed and green lines).

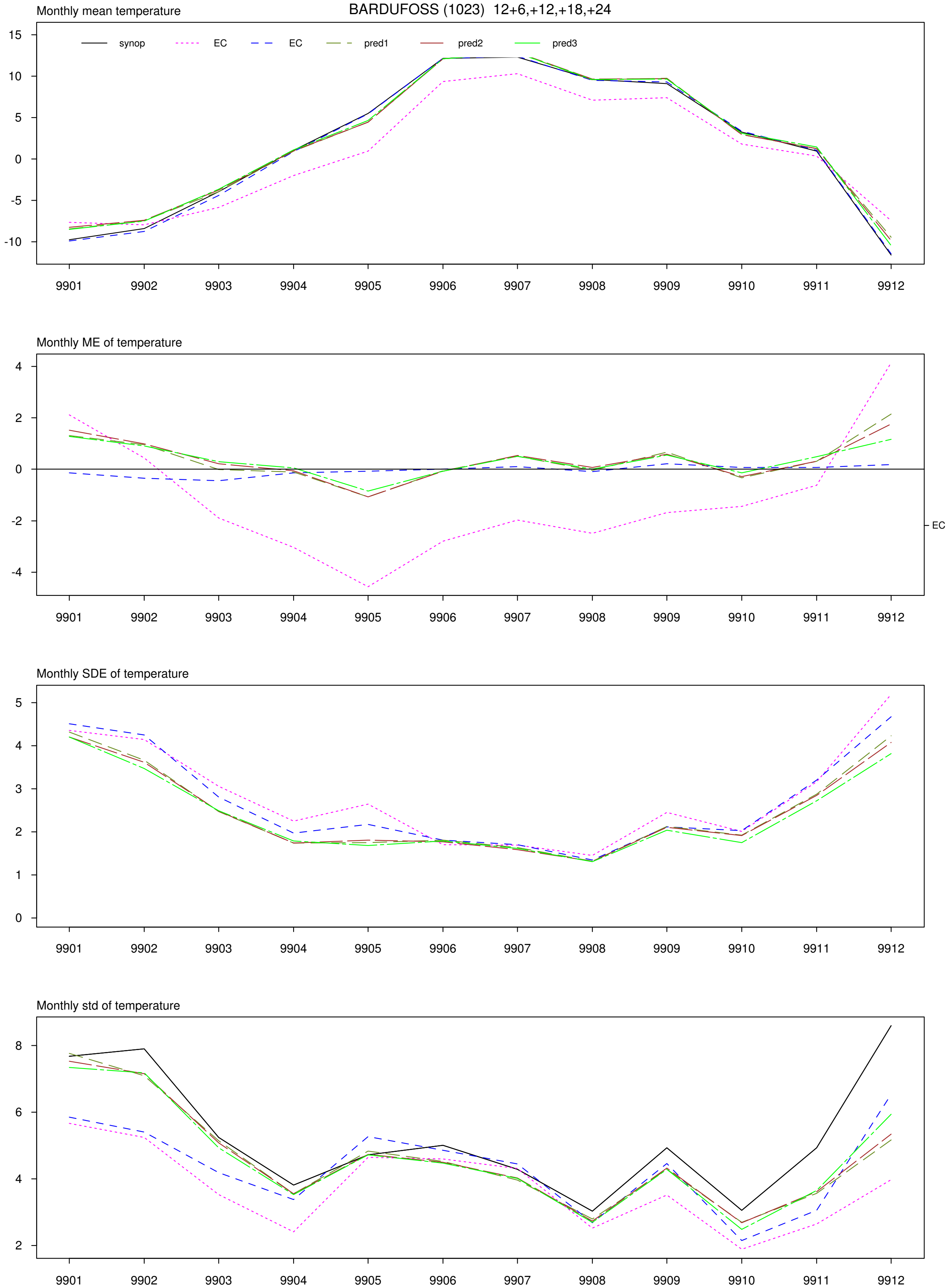


Figure 6: Bardufoss 1999; monthly mean, ME, SDE and standard deviation of temperature observations (black lines) and 12+6,+12,+18,+24 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed), predictions with linear regression based on 1997, 1998 and 2000 data (grey long dashed, red long dashed and green lines).

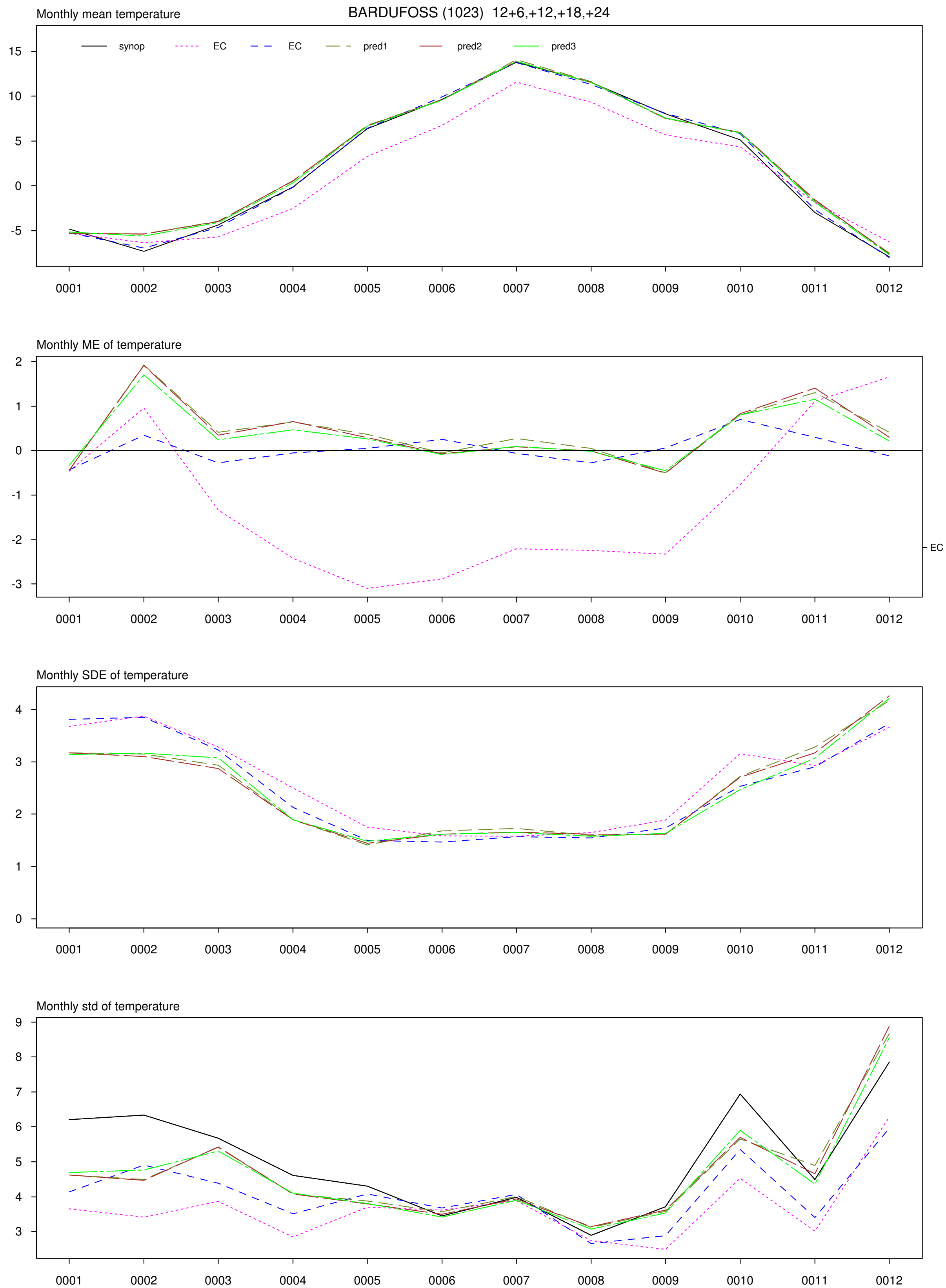


Figure 7: Bardufoss 2000; monthly mean, ME, SDE and standard deviation of temperature observations (black lines) and 12+6,+12,+18,+24 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed), predictions with linear regression based on 1997, 1998 and 1999 data (grey long dashed, red long dashed and green lines).

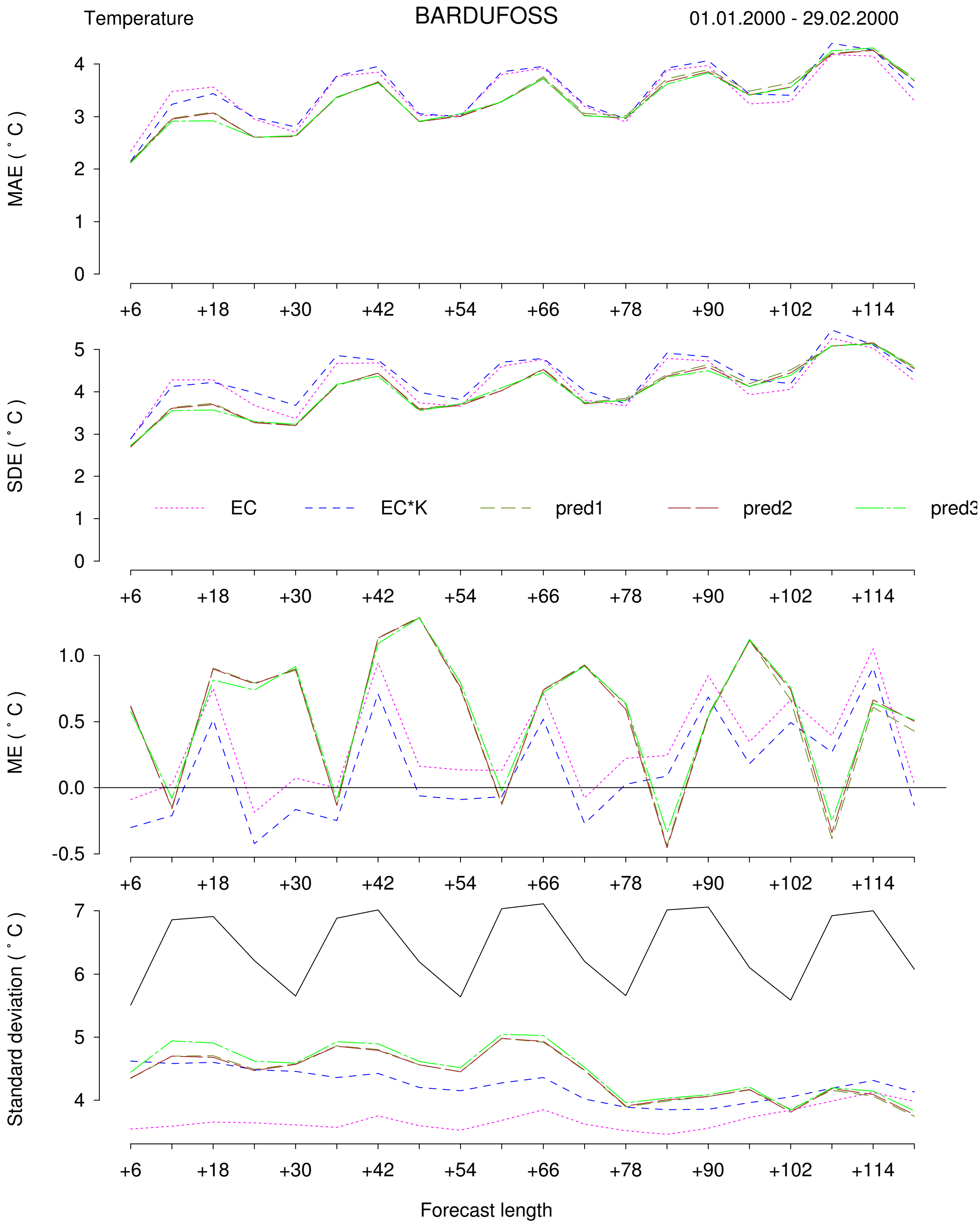


Figure 8: Bardufoss 01.01. - 29.02. 2000; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 12+6,+12, ..., +120 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed), predictions with linear regression based on 1997, 1998 and 1999 data (grey long dashed, red long dashed and green lines).

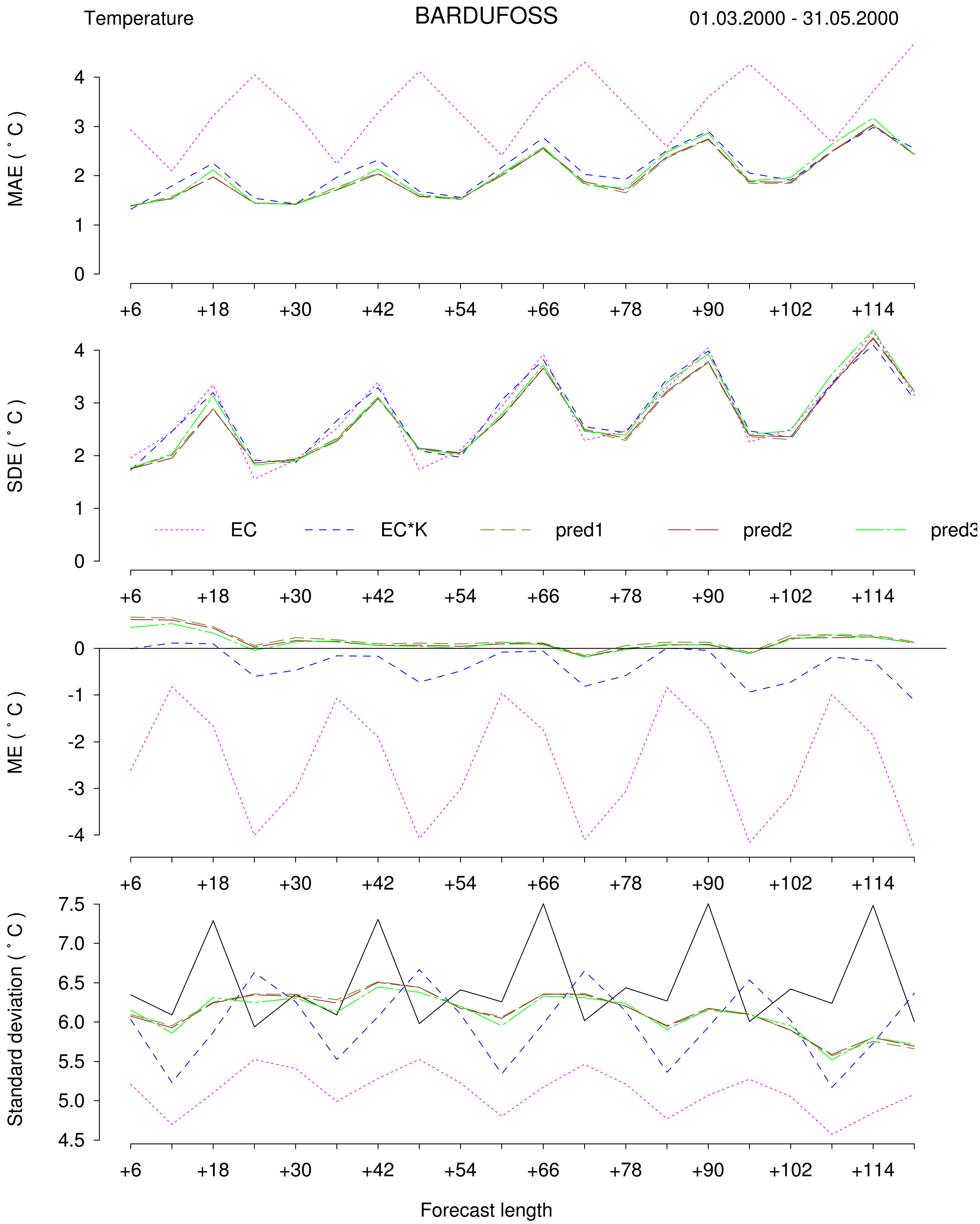


Figure 9: Bardufoss 01.03. - 31.05. 2000; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 12+6,+12, ..., +120 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed), predictions with linear regression based on 1997, 1998 and 1999 data (grey long dashed, red long dashed and green lines).

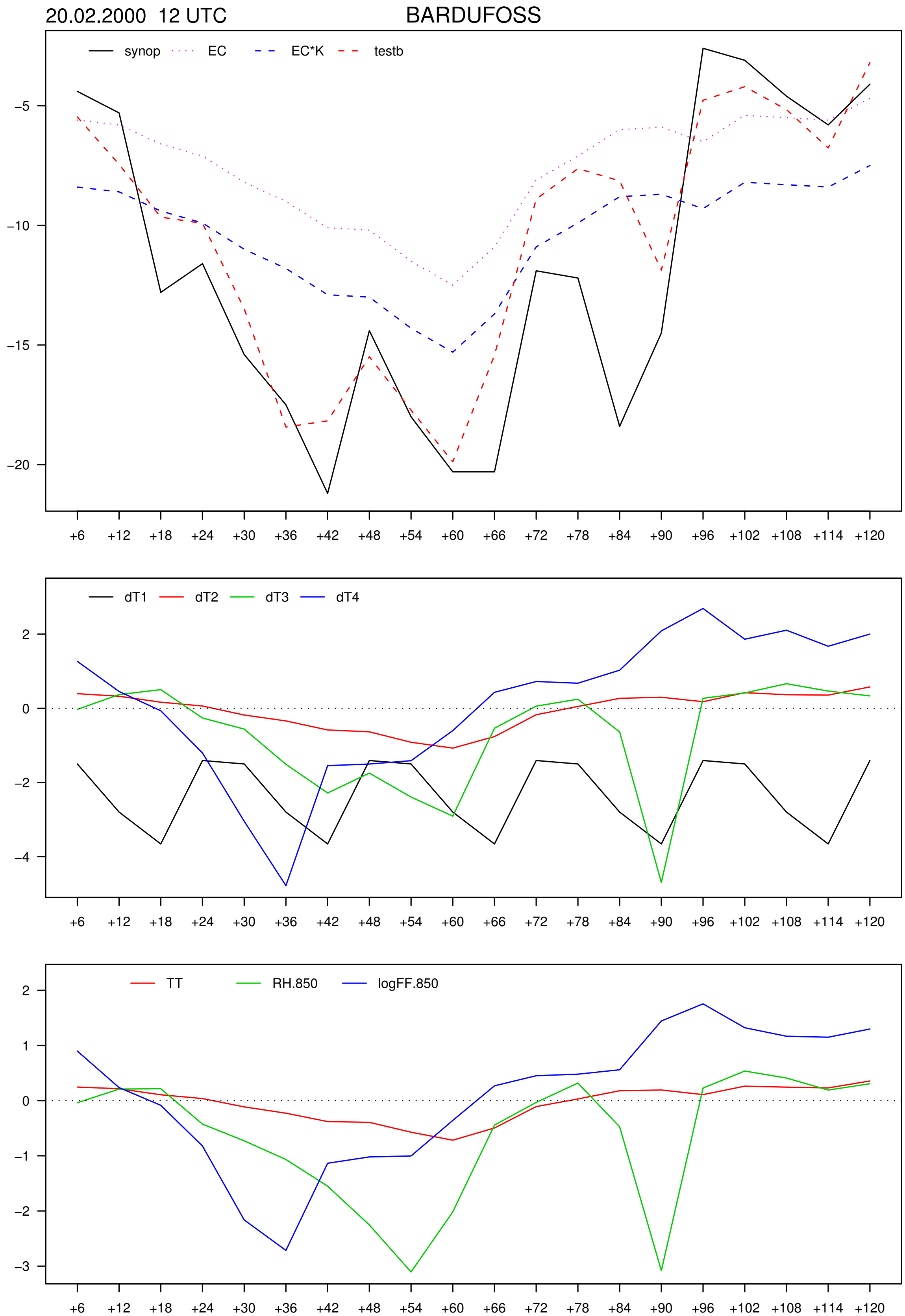


Figure 10: Bardufoss, +6, +12, ..., +120 ECMWF temperature forecasts from 20th of February 2000 12 UTC, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (red dashed), observations (black lines) (top), contributions to the final corrections from the predictors (middle) and normalized predictors (bottom).

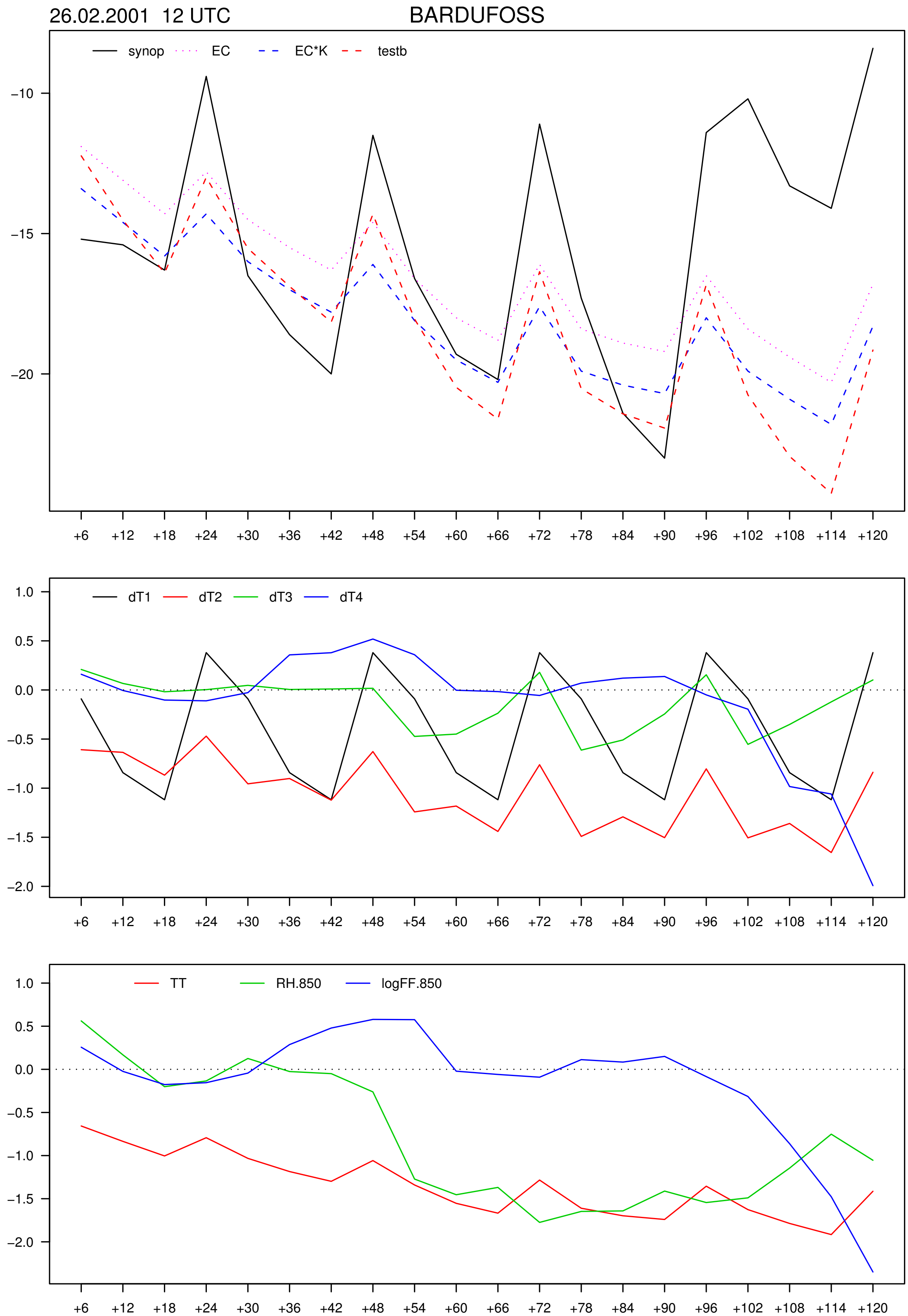


Figure 11: Bardufoss, +6, +12, ..., +120 ECMWF temperature forecasts from 26th of February 2001 12 UTC, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (red dashed), observations (black lines) (top), contributions to the final corrections from the predictors (middle) and normalized predictors (bottom).

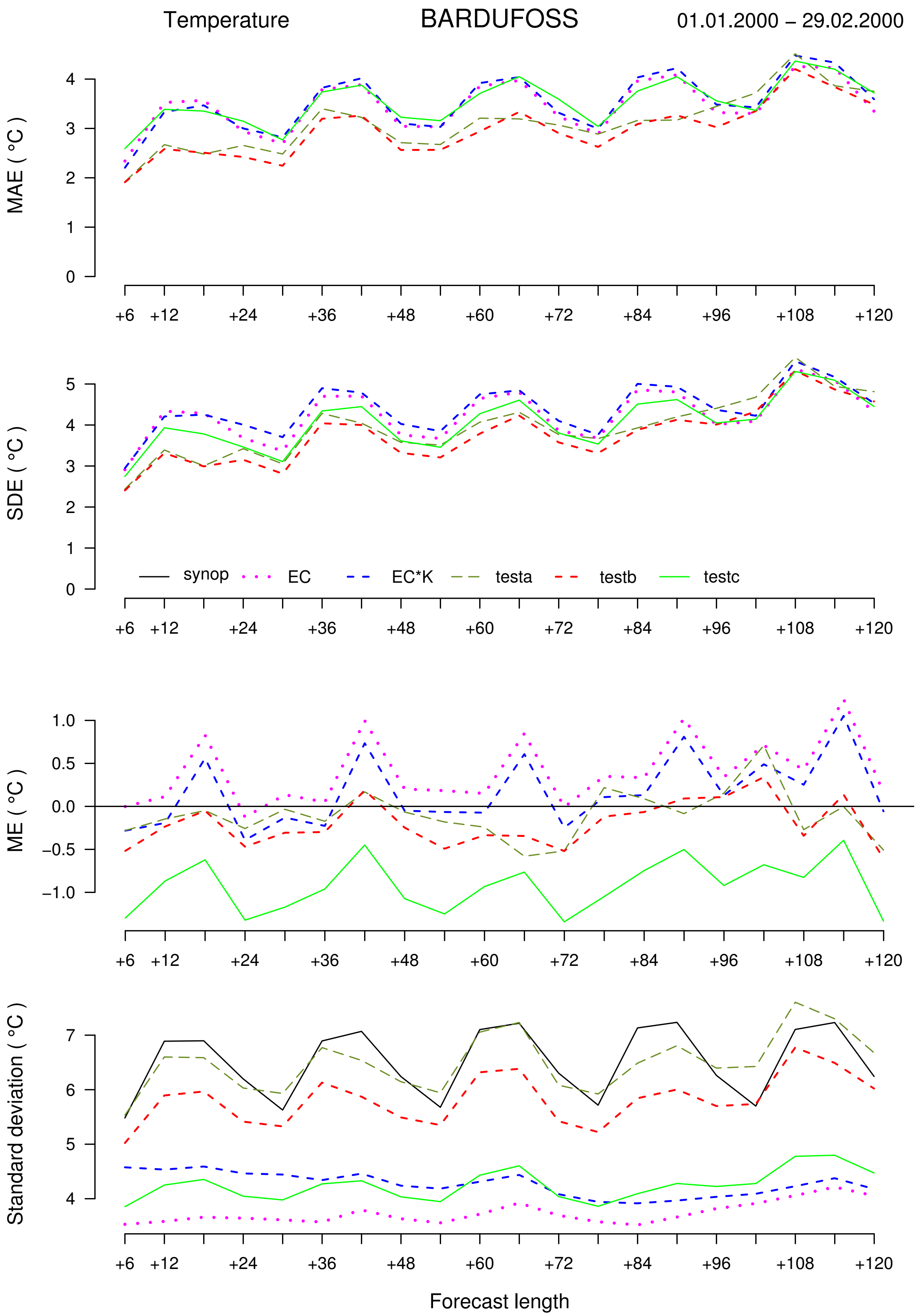


Figure 12: Bardufoss 01.01. - 29.02. 2000; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 12+6,+12, ..., +120 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (grey long dashed, red dashed and green lines).

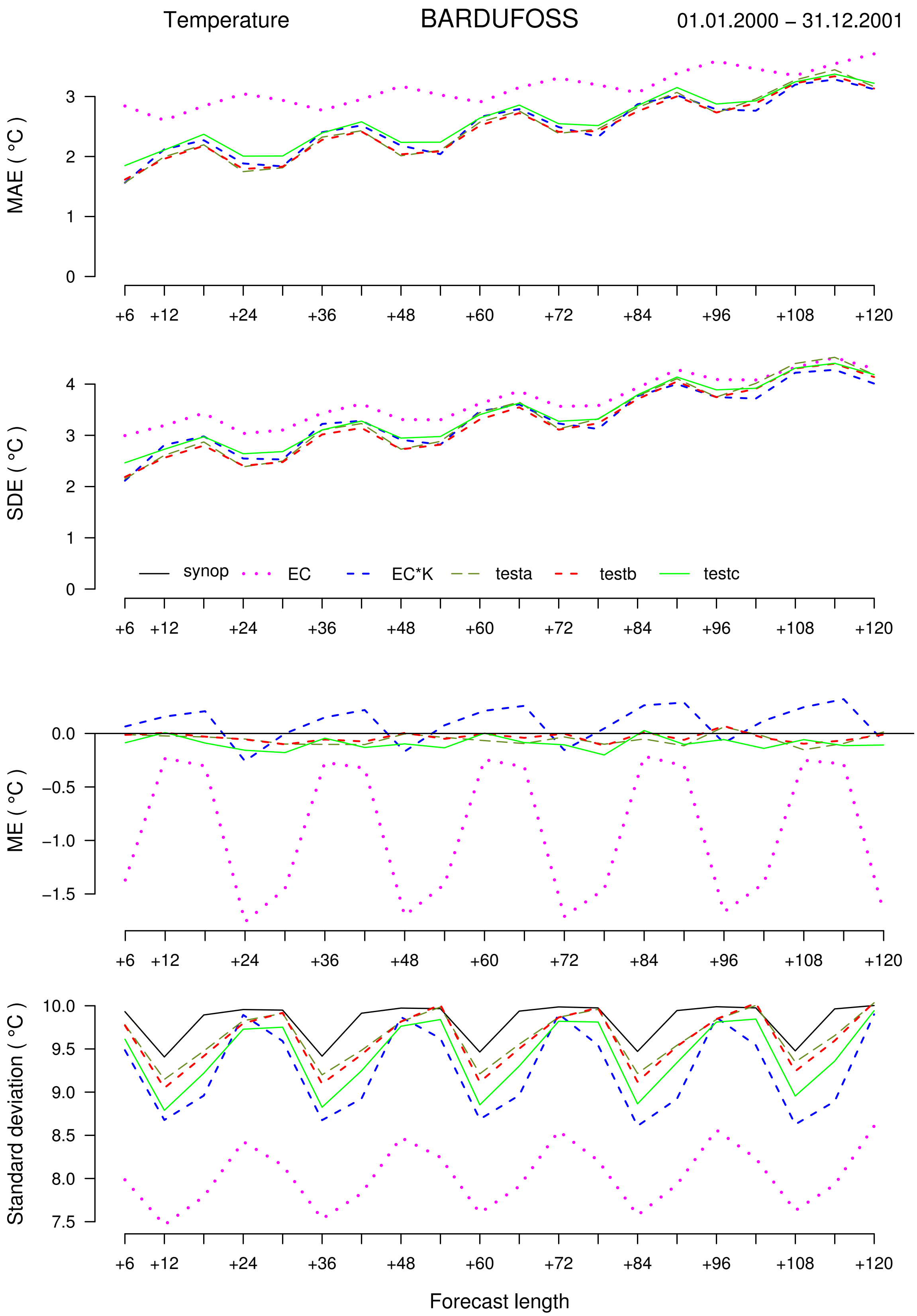


Figure 13: Bardufoss 01.01. 2000 - 31.12. 2001; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 12+6,+12, ..., +120 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (grey long dashed, red dashed and green lines).

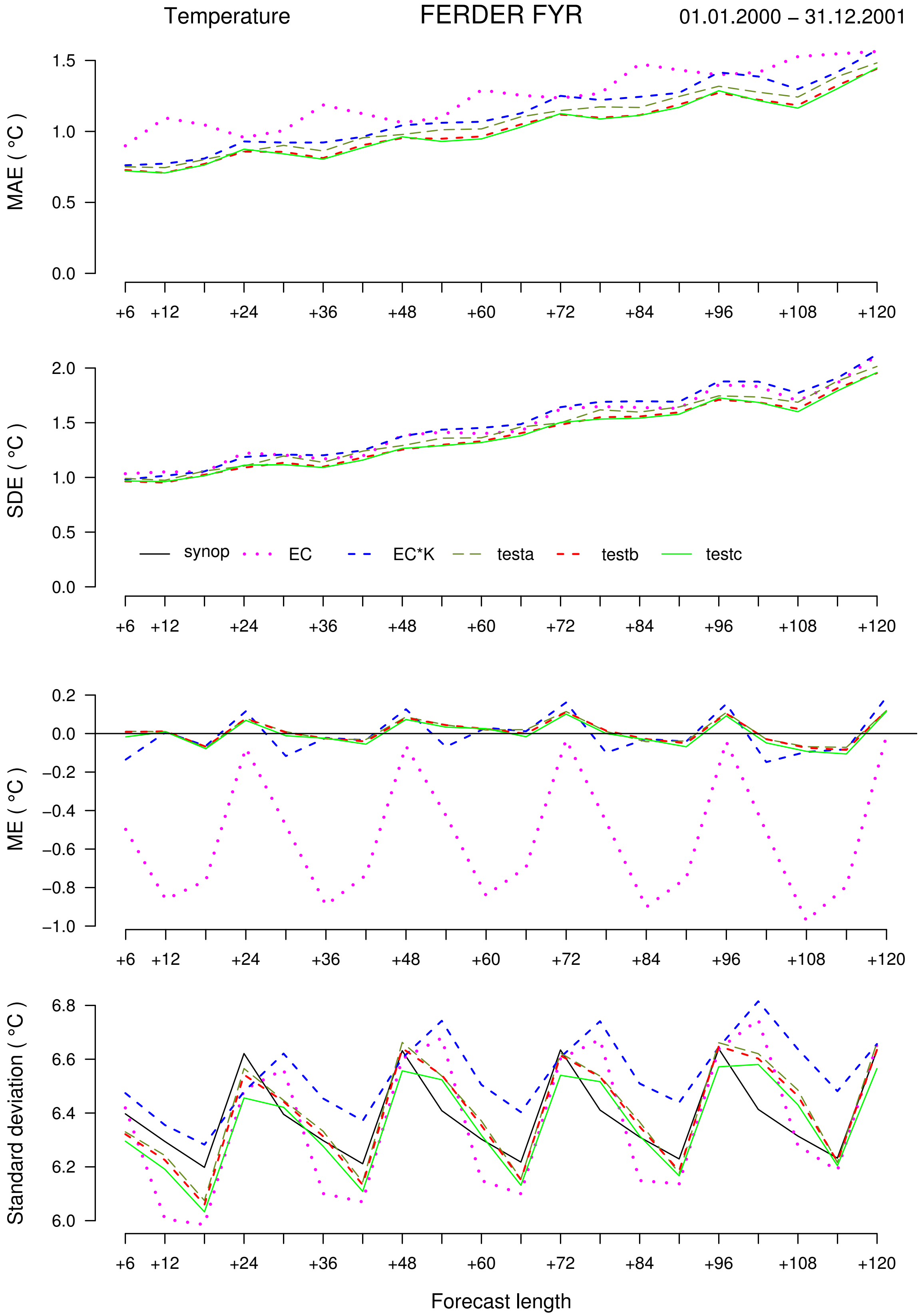


Figure 14: Færder 01.01. 2000 - 31.12. 2001; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 12+6,+12, ..., +120 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (grey long dashed, red dashed and green lines).

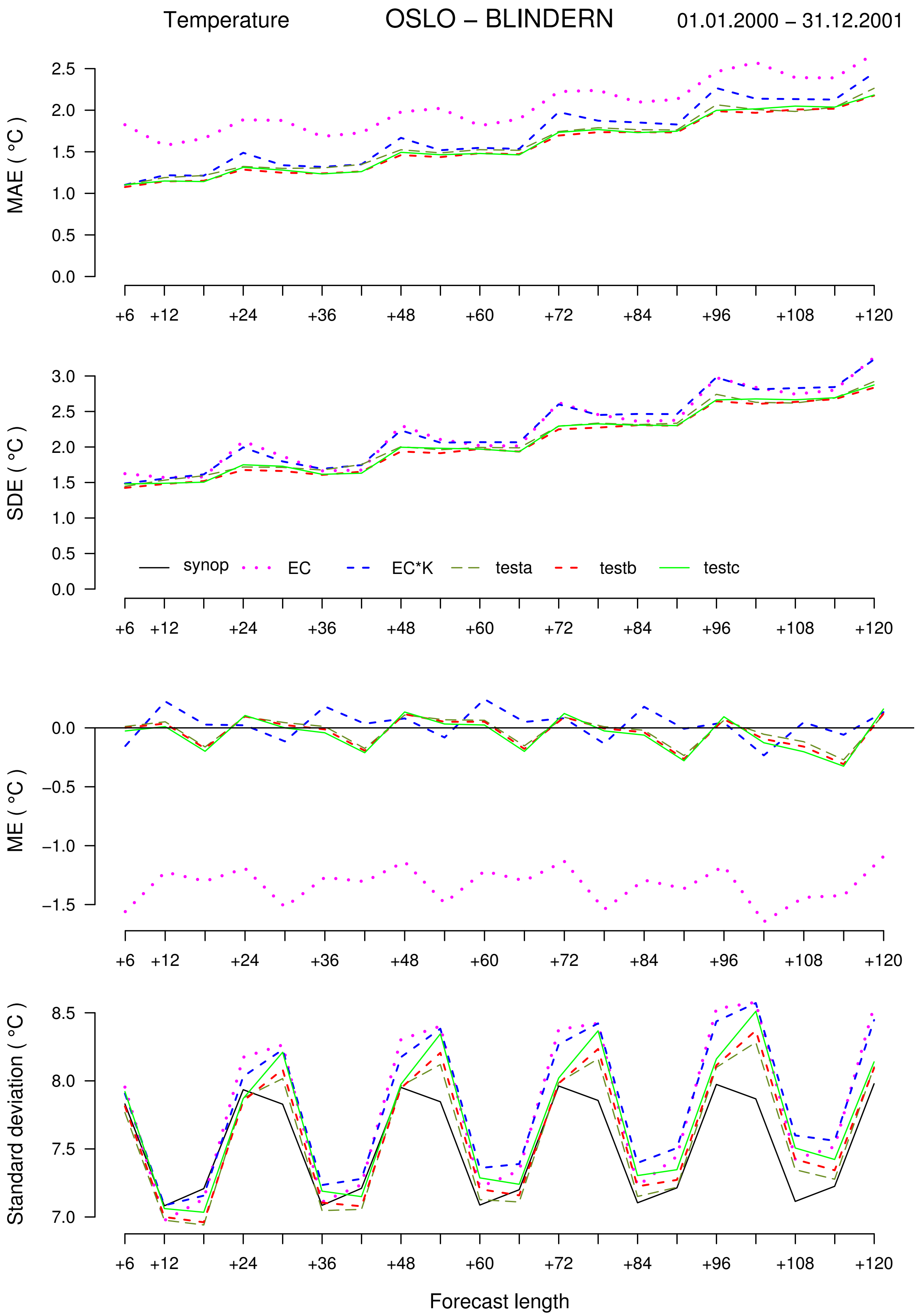


Figure 15: Oslo - Blindern 01.01. 2000 - 31.12. 2001; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 12+6,+12, ..., +120 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (grey long dashed, red dashed and green lines).

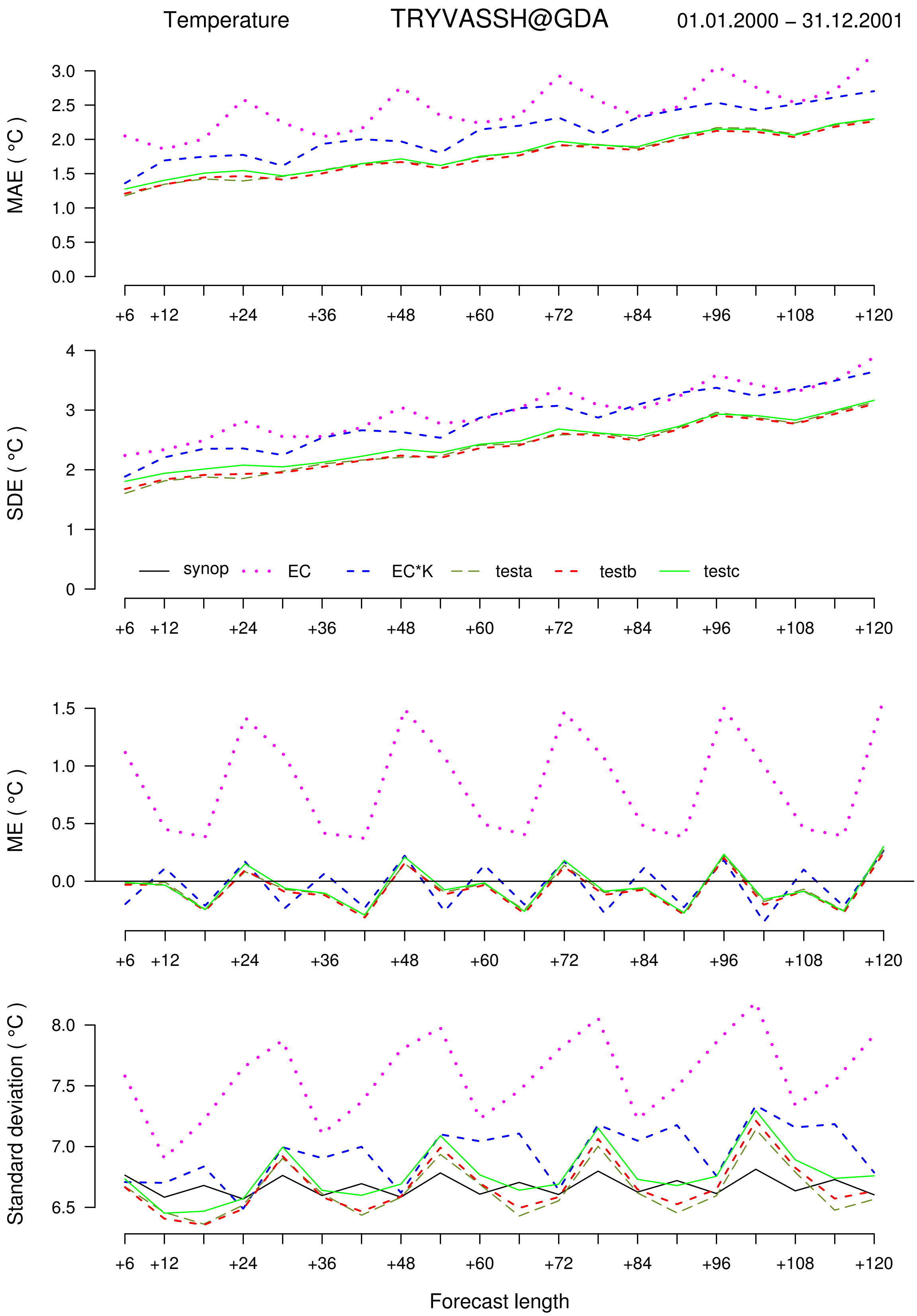


Figure 16: Tryvasshøgda 01.01. 2000 - 31.12. 2001; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 12+6,+12, ..., +120 forecasts from ECMWF, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (grey long dashed, red dashed and green lines).

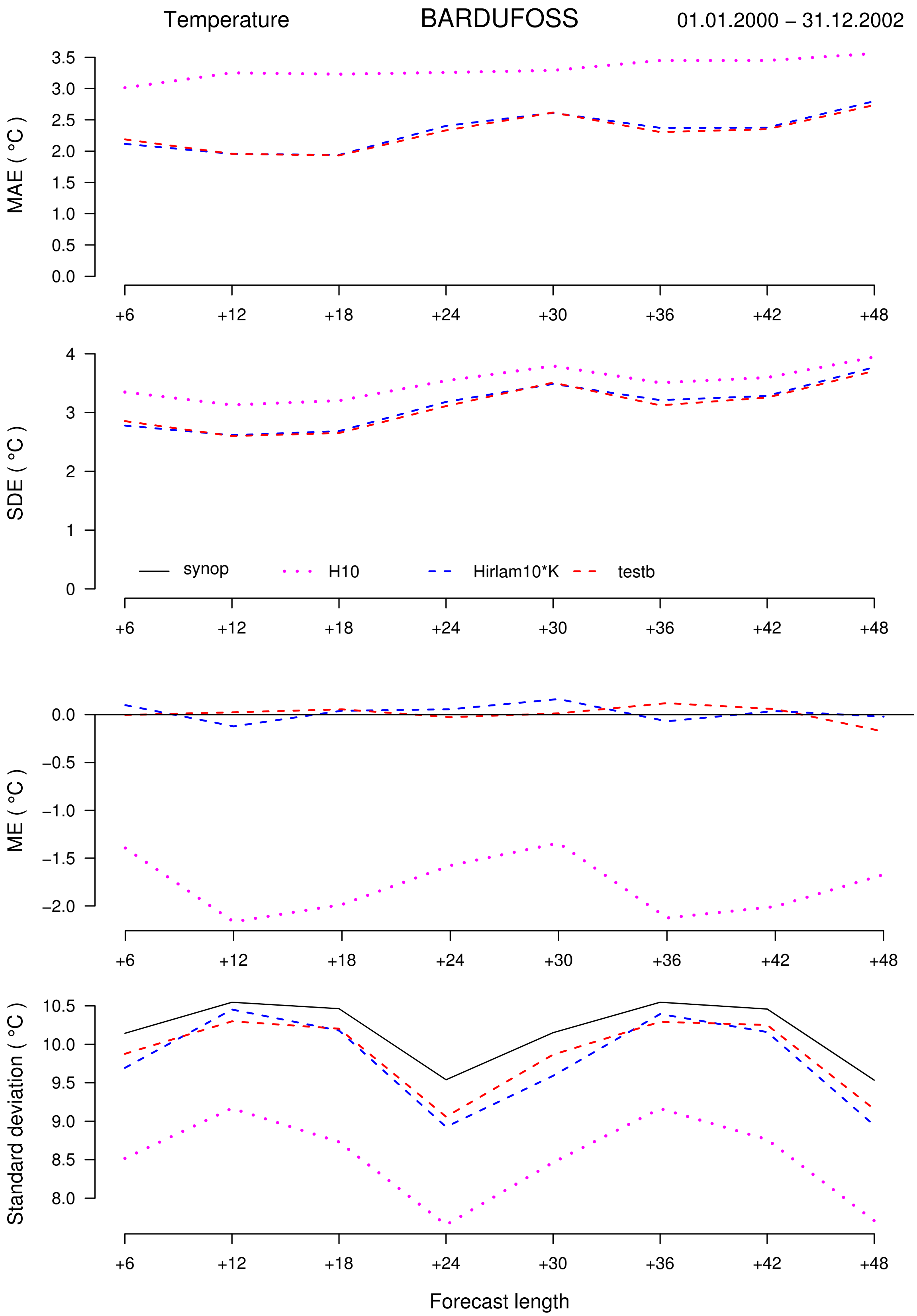


Figure 17: Bardufoss 01.01. 2000 - 31.12. 2002; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 00+6,+12, ... , +48 Hirlam10 forecasts, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (red dashed).

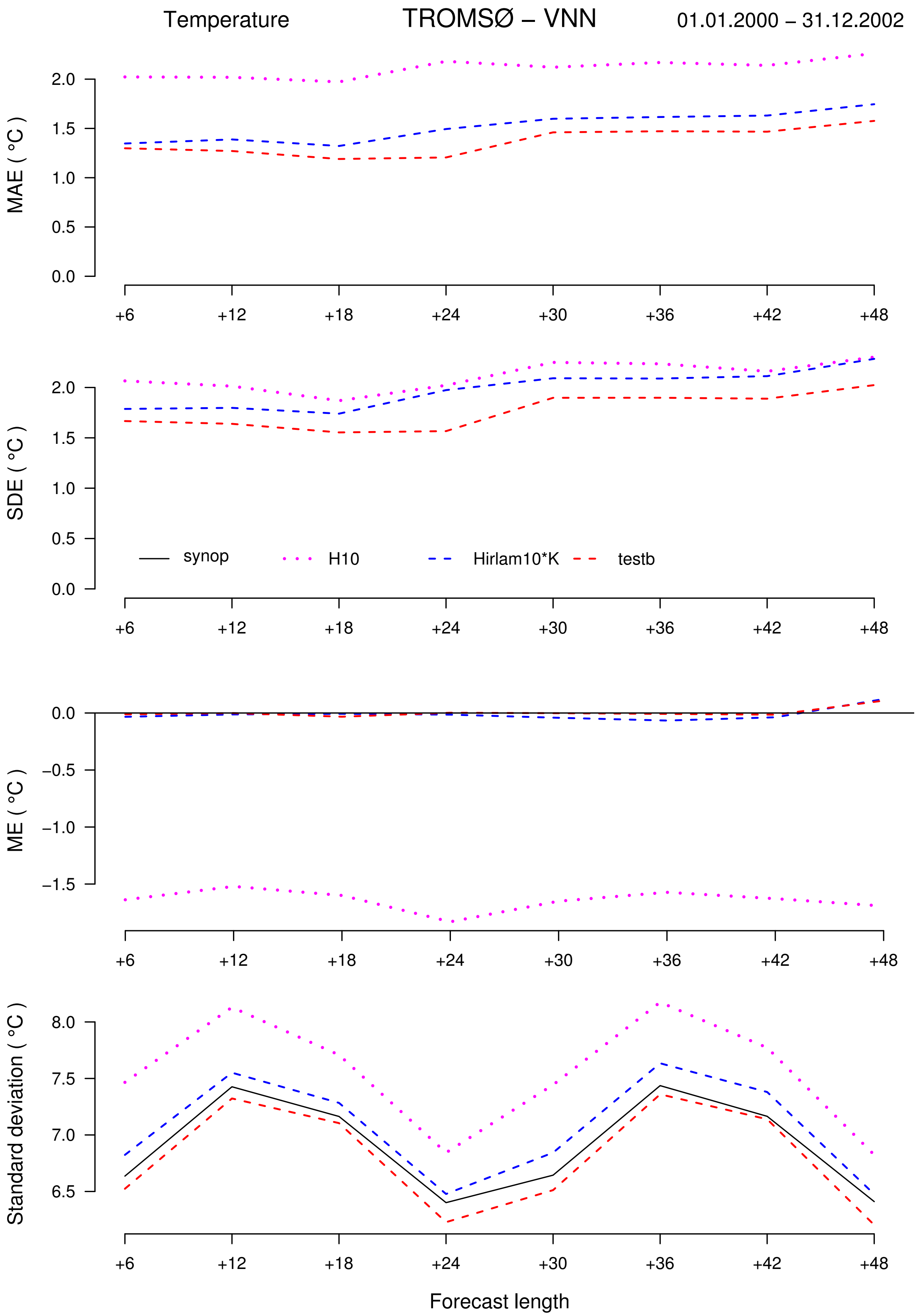


Figure 18: Tromsø 01.01. 2000 - 31.12. 2002; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 00+6,+12, ..., +48 Hirlam10 forecasts, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (red dashed).

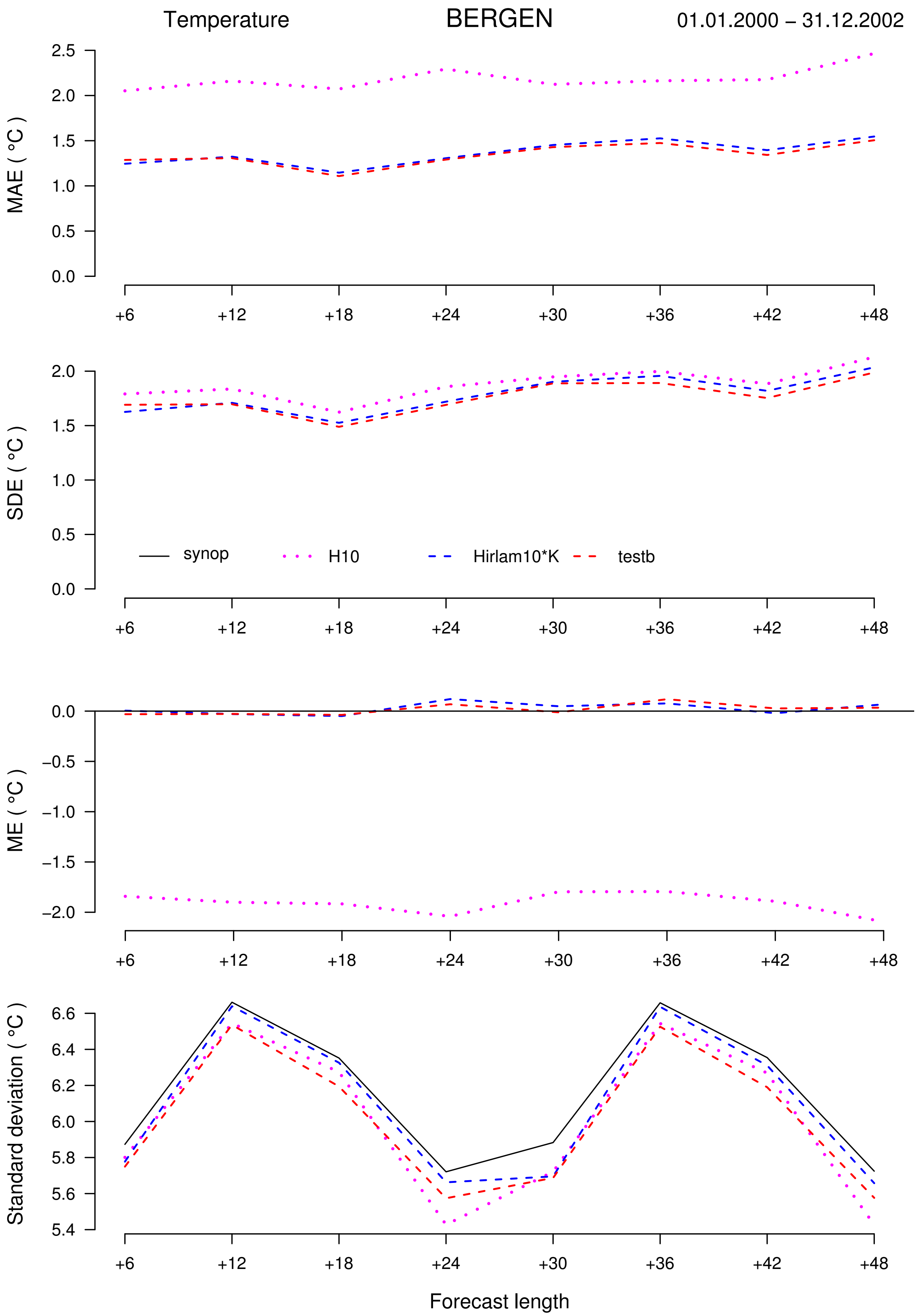


Figure 19: Bergen 01.01. 2000 - 31.12. 2002; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 00+6, 00+12, ..., 00+48 Hirlam10 forecasts, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (red dashed).

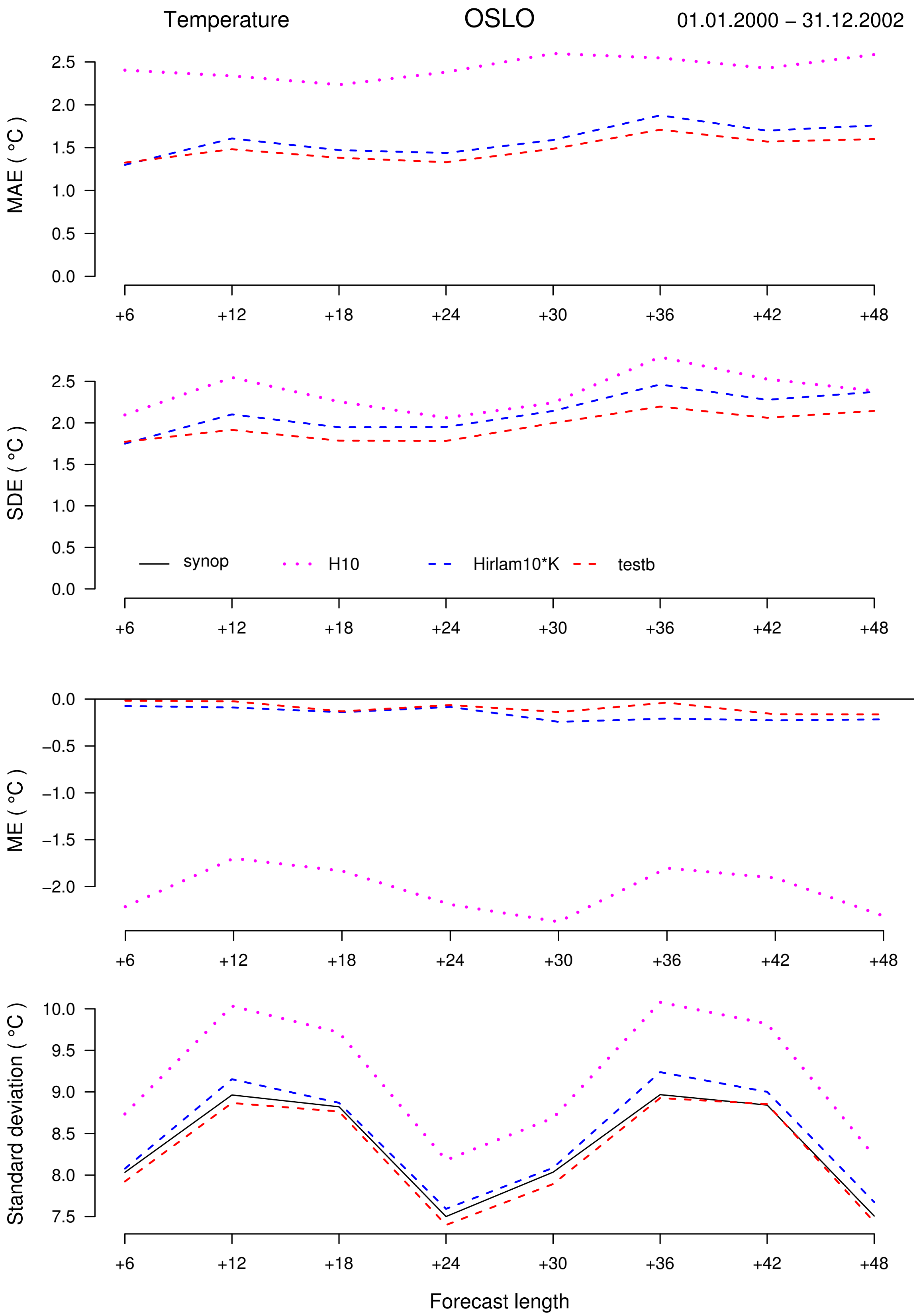


Figure 20: Oslo 01.01. 2000 - 31.12. 2002; MAE, SDE, ME and standard deviation of temperature observations (black lines) and 00+6,+12, ..., +48 Hirlam10 forecasts, uncorrected (pink dotted), corrected by the operational Kalman filter (blue dashed) and the experimental Kalman filter (red dashed).

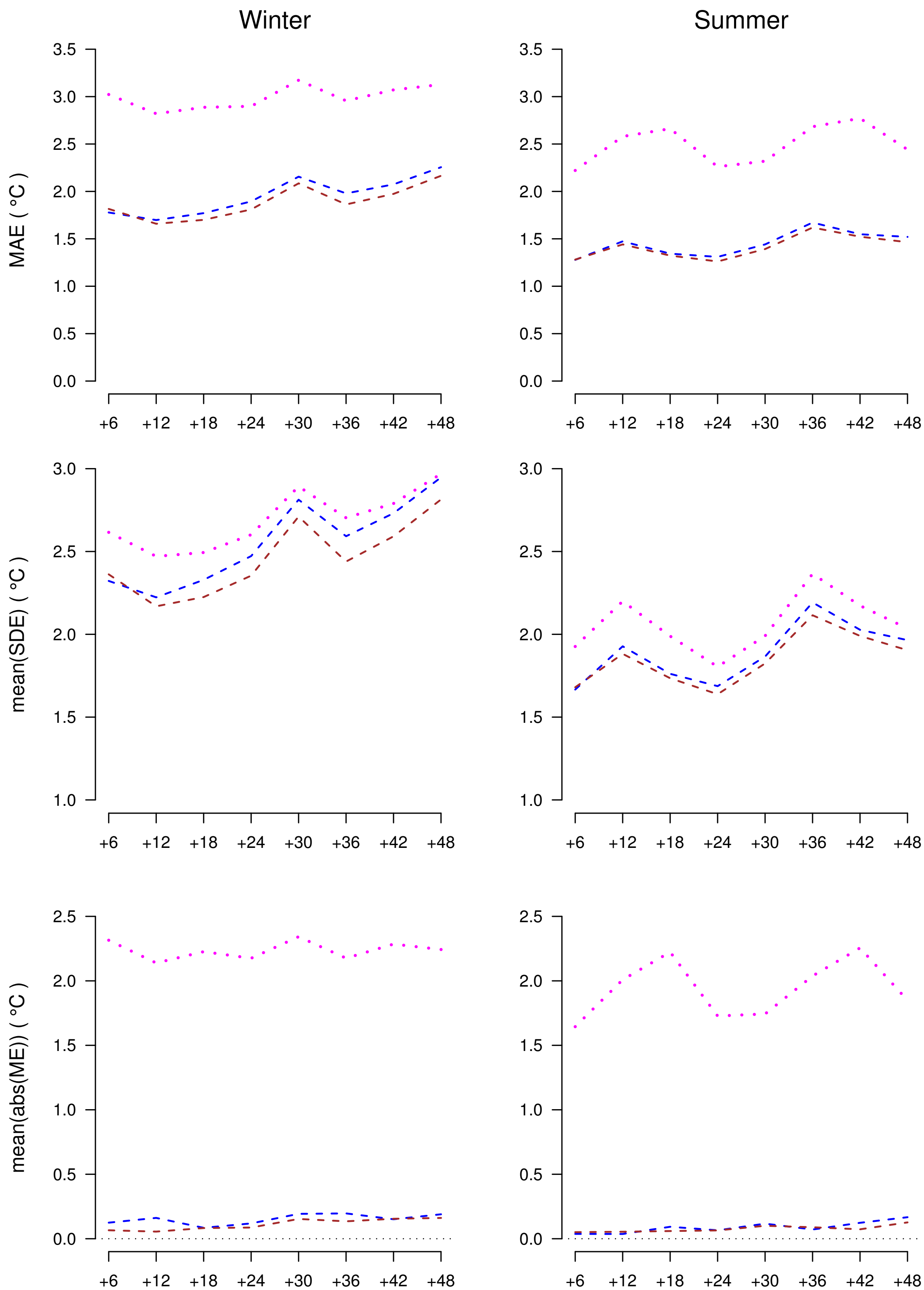


Figure 21: Summary statistics for 37 Norwegian stations for the winter months 2000 - 2002 (left panel) and the summer months (right panel) presented as a function of forecast projection. MAE (upper panel), mean SDE (mid panel) and mean of absolute values of ME are calculated for differences between Hirlam10 forecasts and observations, direct model output (pink dotted lines), operationally corrected (blue dashed lines) and experimentally corrected, testb (red dashed lines).