Meteorol. Appl. 11, 173–187 (2004) DOI:10.1017/S1350482704001252

Adaptive Kalman filtering of 2-metre temperature and 10-metre wind-speed forecasts in Iceland

Philippe Crochet
Icelandic Meteorological Office, Bustadavegur 9, IS-150 Reykjavík, Iceland
Email: philippe@vedur.is

In this paper, an adaptive Kalman filtering procedure is developed and applied to 2-metre temperature and 10-metre wind-speed forecasts in Iceland. The goal is to reduce the systematic bias and improve the accuracy of the local forecasts derived from a numerical weather prediction model and, in addition, to produce reliable prediction intervals. The method consists of adding two algorithms to the traditional Kalman filter procedure that adaptively estimate the noise statistics individually at each forecast location and each time step. The tested data set shows that the method is able to remove the systematic errors and to quantify the prediction uncertainty in a consistent manner.

1. Introduction

Statistical adaptations of numerical weather prediction (NWP) model outputs are often necessary to cope with local conditions and adjust the information delivered at a coarse spatial resolution to a specific location. Two common statistical methods that are used to locally adjust the forecasts of surface parameters such as the 2-metre temperature (T_{2m}) and the 10-metre wind-speed (FF_{10}) are the Model Output Statistics (MOS) method (Glahn & Lowry 1972) and the Perfect Prog Method (PPM) (Klein & Lewis 1970). These methods are based on the development of regression equations between the observed parameter to be predicted and a set of explanatory variables that are either forecasted (MOS), or either observed or analysed (PPM). The development of such equations usually requires an extensive data set of historical observations. In the case of MOS, the validity of the equations may become questionable if the NWP model characteristics are modified and a new calibration becomes necessary. This recalibration can be performed, for instance, within the framework of the updatable MOS, UMOS (Wilson & Vallée 2002, 2003). The Kalman filter (Kalman 1960) is an adaptive method which offers an alternative solution to the standard regression models and does not suffer from these drawbacks. It is defined by a set of recursive relationships which combine current measurements and forecasts of the variable under study in order to infer the statistical properties of future observations of this variable. Applications in weather forecasting can be found in Golanis & Anadranistokis (2002), Homleid (1995), Persson (1991) and Simonsen (1991). The Kalman filter (KF) procedure can be of variable complexity, ranging from the simple bias adjustment to more sophisticated and flexible modelling involving one or more predictors. Since 1996, the T_{2m} and FF_{10} local forecasts derived from the European Centre for Medium Range Weather Forecasts (ECMWF) deterministic T511 NWP model, 1200 UTC run (EC-12) have been post-processed at the Icelandic Meteorological Office (IMO) by applying a KF procedure. The algorithm used so far at IMO is with non-adapted noise statistics and makes use of a set of predefined values for the observation noise variance V_t and the system noise variance W_t (where t is a time index). In the present paper, a modification of the current operational KF algorithm used at IMO is studied. The goal is to reduce the systematic forecasting errors, to better regulate the capacity of adaptation of the algorithm when large errors occur and to produce reliable prediction intervals. The modification consists of extending the concept of adaptivity to the noise statistics as well. This is performed by adding two online algorithms that sequentially estimate V_t and W_t, at each time step, individually for each forecast range and each location on the basis of the previous forecasting errors. Section 2 presents briefly the adaptive KF algorithm, section 3 indicates the potential of the method and section 4 concludes this paper.

2. The adaptive Kalman filter

Below is a brief presentation of the principle of the adaptive KF procedure.

Let $Y_t, Y_{t-\Delta t}, \ldots, Y_{t-k\Delta t}, \ldots, Y_{t_0}$ denote the observed values of the process under study (T_{2m} or FF_{10}) at equally spaced time intervals Δt: $t, t-\Delta t, \ldots, t-k\Delta t, \ldots, t_0$. Y_t is assumed to depend on a $n \times 1$ state vector X_t and a known $1 \times n$ vector H_t of predictors.

The state vector X_t is linearly related to Y_t through the observation equation:

$$Y_t = H_t X_t + v_t \quad (1)$$

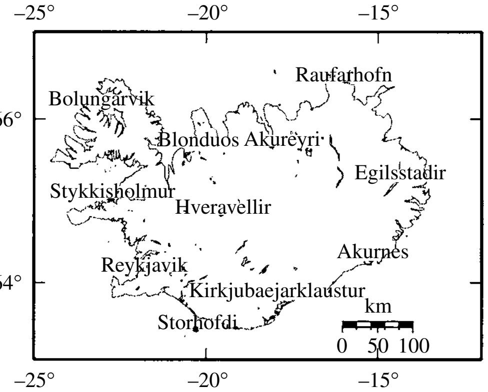

Figure 1. *The 11 sites of the local forecasts over Iceland.*

where v_t is the observation noise and is assumed to be normally distributed with mean zero and variance V_t, $v_t \sim N(0, V_t)$. By comparison with a linear regression, H_t represents the predictors or explanatory variables and the state vector X_t gives the regression coefficients which are here allowed to vary at each time step. This time dependent evolution of X_t is defined by the system equation:

$$X_t = \Phi X_{t-k\Delta t} + \Gamma w_t \tag{2}$$

where Φ is a known $n \times n$ transition matrix, Γ a $n \times r$ matrix of fixed coefficients, and w_t a $r \times 1$ vector representing the system noise which is assumed to be normally distributed with mean zero and variance W_t, $w_t \sim N(0, W_t)$, and independent of v_t. The observation and system equations define a dynamic linear model (DLM) (see Harrison & Stevens 1976). The observation equation describes the stochastic dependence of the observations on the process parameters and the system equation describes the time evolution of the process parameters. The KF algorithm is a recursive relationship used to identify the parameters of the DLM through a straightforward application of the Bayes theorem (see Meinhold & Singpurwalla 1983). Detailed description of the algorithm is given in Appendix 1. The design of the KF procedure requires the knowledge of V_t and W_t. Homleid (1995) suggests defining these parameters by using a statistical estimation procedure or by tuning them in order to make the KF behave as requested, for instance to allow the KF to react quickly to new conditions. Simonsen (1991) proposes giving an estimate according to the system designer expectations. Persson (1991) suggests manipulating the covariances through external interference with respect to external conditions. In a number of practical situations, the estimation of these parameters may prove to be difficult for the modeller, when for instance a large number of locations are considered with various climatic and/or geographic conditions. To help with this, several online methods can be used to adaptively estimate V_t and W_t individually for each location in an automatic manner. Golanis & Anadranistokis (2002) define an estimation procedure for W_t and V_t based on the sample of the last 7 values. West et al. (1985) suggest the use of a discounting method for W_t and a learning procedure for V_t. Prior to that work, Mehra (1970) described rigorous methods for estimating V_t and W_t. In the present work, two computationally simple methods are adopted for estimating V_t and W_t. The observation noise variance V_t is sequentially estimated with the Smith algorithm (Smith 1967), described in Appendix 2. The system noise variance, W_t, is estimated with the Jazwinski algorithm (Jazwinski 1969), given in Appendix 3. This algorithm takes care of the problem of filter divergence and stiffness observed when the KF becomes overconfident and the impact of incoming observations is very limited. The Jazwinski algorithm detects such behaviour and action is taken in order to modify the sensitivity of the filter by inflating the system noise variance in a suitable manner. Both algorithms used

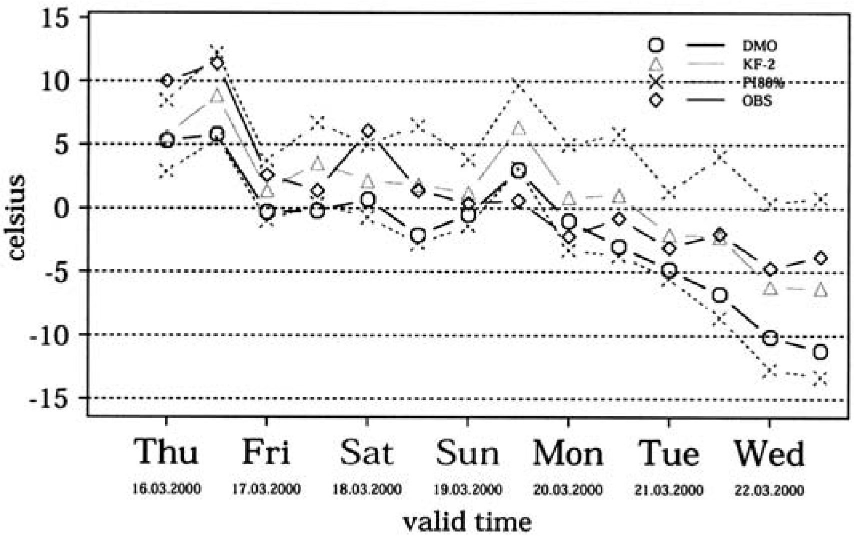

Figure 2. *Two-metre temperature forecast for Akureyri, 15 March 2000 at 1200 UTC, using the KF-2 protocol. The DMO forecast, the observed temperature and the 80% prediction interval (PI) derived from the KF procedure are also represented. For all the forecasts valid at 0000 UTC, the adjustment is as follows: KF-2 = DMO*0.766 + 1.61. For all the forecasts valid at 1200 UTC, the adjustment is as follows: KF-2 = DMO*0.89 + 3.72.*

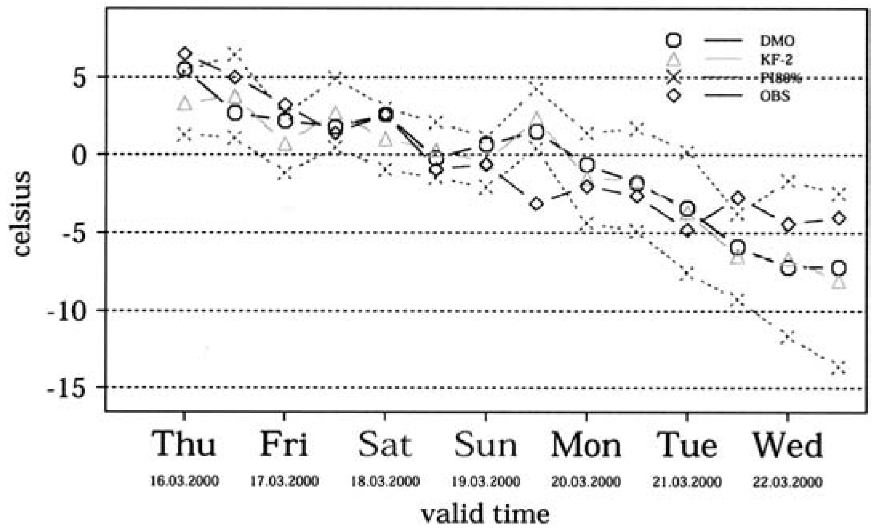

Figure 3. *As Figure 2 but for Stykkisholmur. For all the forecasts valid at 0000 UTC, the adjustment is as follows: KF-2 = DMO*0.78 -1.0. For all the forecasts valid at 1200 UTC, the adjustment is as follows: KF-2 = DMO*1.19 + 0.56.*

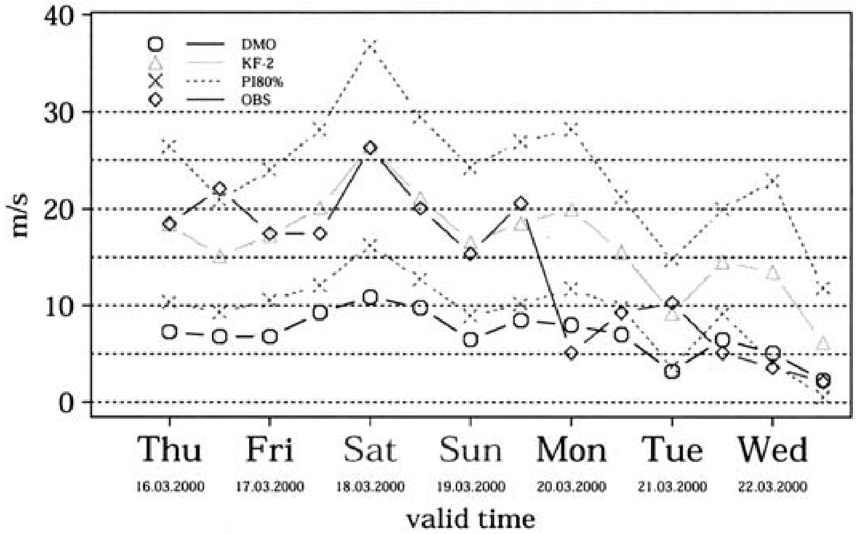

Figure 4. *Ten-metre wind-speed forecast for Hveravellir, 15 March 2000 at 1200 UTC, using the KF-2 protocol. The DMO forecast, the observed temperature and the 80% prediction interval (PI) derived from the KF procedure are also represented. For all the forecasts valid at 0000 UTC, the adjustment is as follows: KF-2 = DMO*2.47 + 2.42. For all the forecasts valid at 1200 UTC, the adjustment is as follows: KF-2 = DMO*1.99 + 1.60.*

here are described in Sage & Husa (1969) who review a large set of existing methods for the online adaptive estimation of noise statistics. Once the adaptive KF algorithm is set up, it enables inference to be made about the statistical properties of future observations, namely the mean ($\hat{\mu}_{t|t-k\Delta t}$) and variance ($\hat{\sigma}^2_{t|t-k\Delta t}$) of the predictive distribution of Y_t made at time $t - k\Delta t$ (see Eq. A-4 and A-5 in Appendix 1). From this information, the $100(1-\alpha)\%$ prediction interval for Y_t can be defined as follows:

$$\left(\hat{\mu}_{t|t-k\Delta t} - U_{\alpha/2}\hat{\sigma}_{t|t-k\Delta t}, \hat{\mu}_{t|t-k\Delta t} + U_{\alpha/2}\hat{\sigma}_{t|t-k\Delta t}\right) \quad (3)$$

where $U_{\alpha/2}$ denotes the upper $\alpha/2$ quantile of the standard normal distribution and $\hat{\sigma}_{t|t-k\Delta t}$ the standard-deviation of the predictive distribution of Y_t made at time $t - k\Delta t$. This prediction interval quantifies the reliability of the local prediction.

3. Application of the adaptive KF procedure in weather forecasting

In this section, the adaptive KF algorithm described above is applied to the predictions of T_{2m} and FF_{10} in Iceland.

3.1. The data

The adaptive KF algorithm was run over the period 1 January 2000–1 March 2001 for post-processing local DMO forecasts of T_{2m} and FF_{10} derived from EC-12, at locations corresponding to 11 synoptic stations in Iceland (Figure 1). For three of these stations (Blonduos, Bolungarvik and Kirkjubaejarklaustur), the wind measurements are made visually, using qualitative

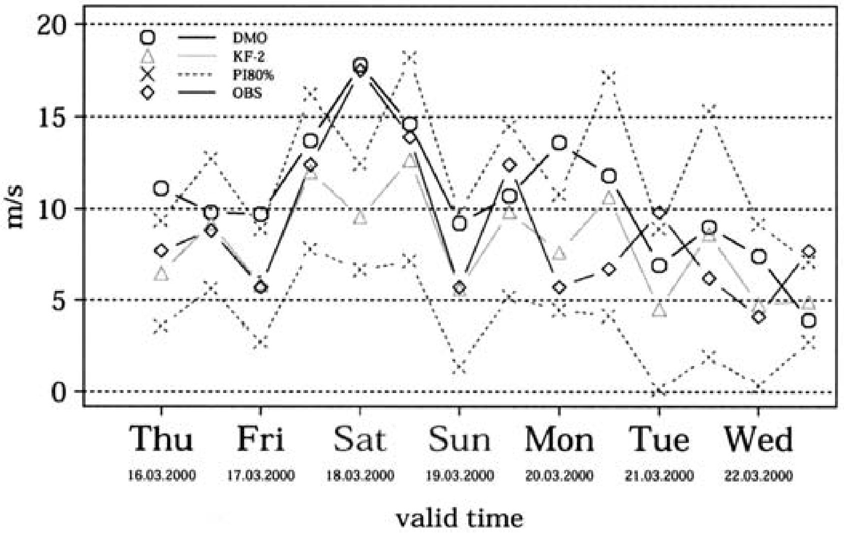

Figure 5. *As Figure 4 but for Reykjavik. For all the forecasts valid at 0000 UTC, the adjustment is as follows: KF-2 = DMO*0.46 + 1.32. For all the forecasts valid at 1200 UTC, the adjustment is as follows: KF-2 = DMO*0.72 + 2.08.*

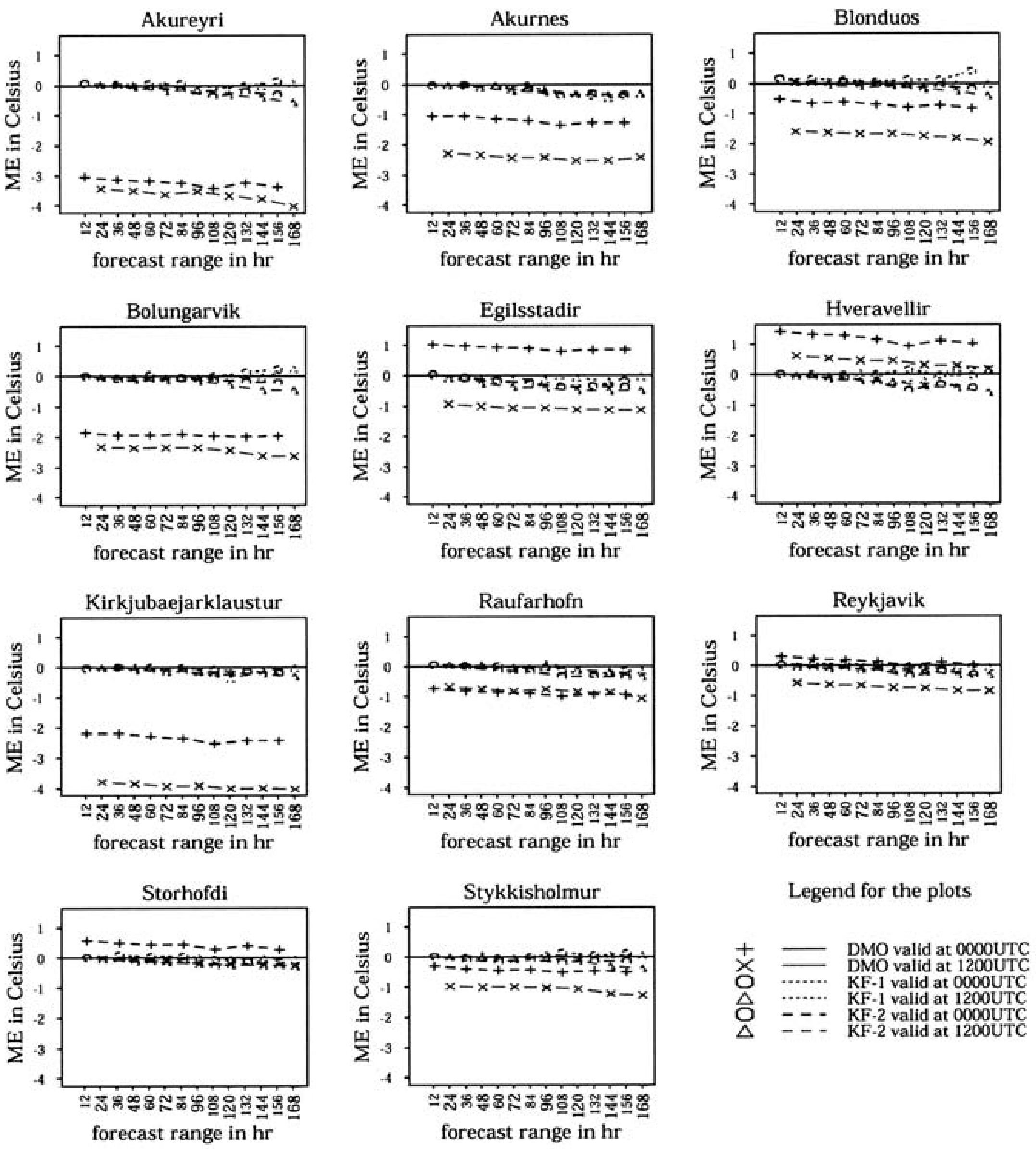

Figure 6. *Mean error (ME) for the 2-metre temperature forecasts, calculated for each location over the period 1 March 2000–1 March 2001 and presented as a function of the forecast range.*

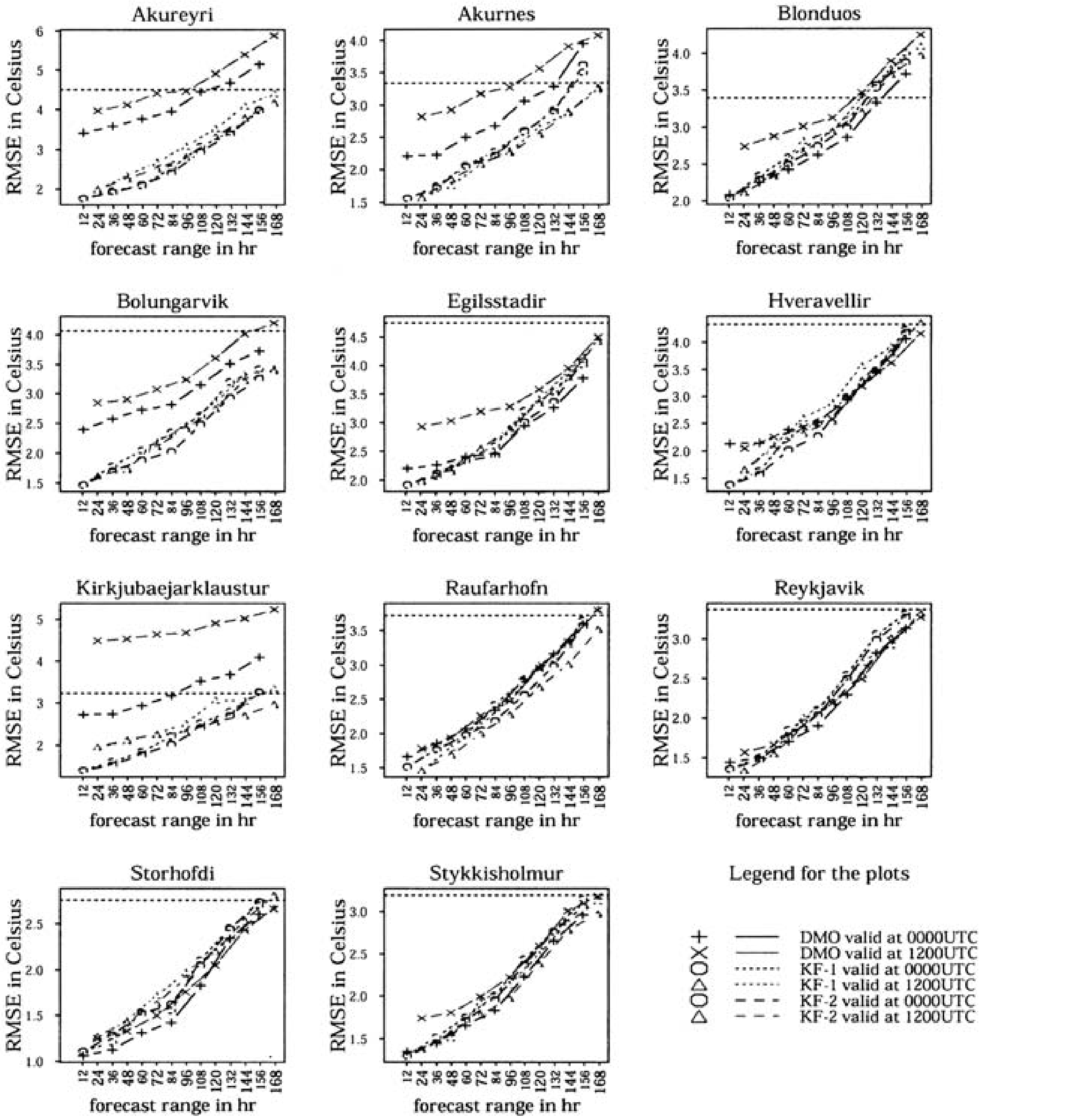

Figure 7. *Root mean squared error (RMSE) for the 2-metre temperature forecasts, calculated for each location over the period 1 March 2000–1 March 2001 and presented as a function of the forecast range. The horizontal dashed line gives the RMSE for a climatic prediction.*

signs according to the Beaufort scale of wind (WMO 1988). The EC-12 NWP model is run once every 24 hours using initial information from the analysis made at 1200 UTC. The spatial resolution of the DMO data used in this study is 1.5°. These DMO data are then interpolated to each location using a bilinear interpolation of the four nearest grid points. The forecast projection times (KF time lags) considered here range from + 12 hr to + 168 hr in 12 hr steps (Δt). The observations made at 0000 and 1200 UTC are used to update the parameters of the KF procedure.

3.2. Experimental protocol

Two experimental protocols are compared in this study:

1. KF-1: the KF algorithm is applied as described in Appendix 1. Each forecast issued at time t and valid at a given projection time is treated independently of the other projection times. A forecast issued at time t (1200 UTC) and valid at time $t + k\Delta t$ (1200 UTC), i.e. $k = 2,4,6,8,10,12,14$, is updated with information from a forecast of the same forecast range, i.e. issued at time $t - k\Delta t$ (1200 UTC) and valid at time t (1200 UTC). A forecast issued at time t (1200 UTC) and valid at time $t + k\Delta t$ (0000 UTC), i.e. $k = 1,3,5,7,9,11,13$, is updated with information from a forecast issued at time t-$(k+1)\Delta t$ (1200 UTC) and valid at time $t - \Delta t$ (0000 UTC). This means that there are $k\Delta t/24$ different KF series for a given forecast range $k\Delta t$ valid at 1200 UTC and $(k+1)\Delta t/24$ different KF series for a given forecast range $k\Delta t$ valid at 0000 UTC.
2. KF-2: a certain linear dependency between the batch of forecasts issued at instant t and valid at the same verification time (either 0000 or 1200 UTC)

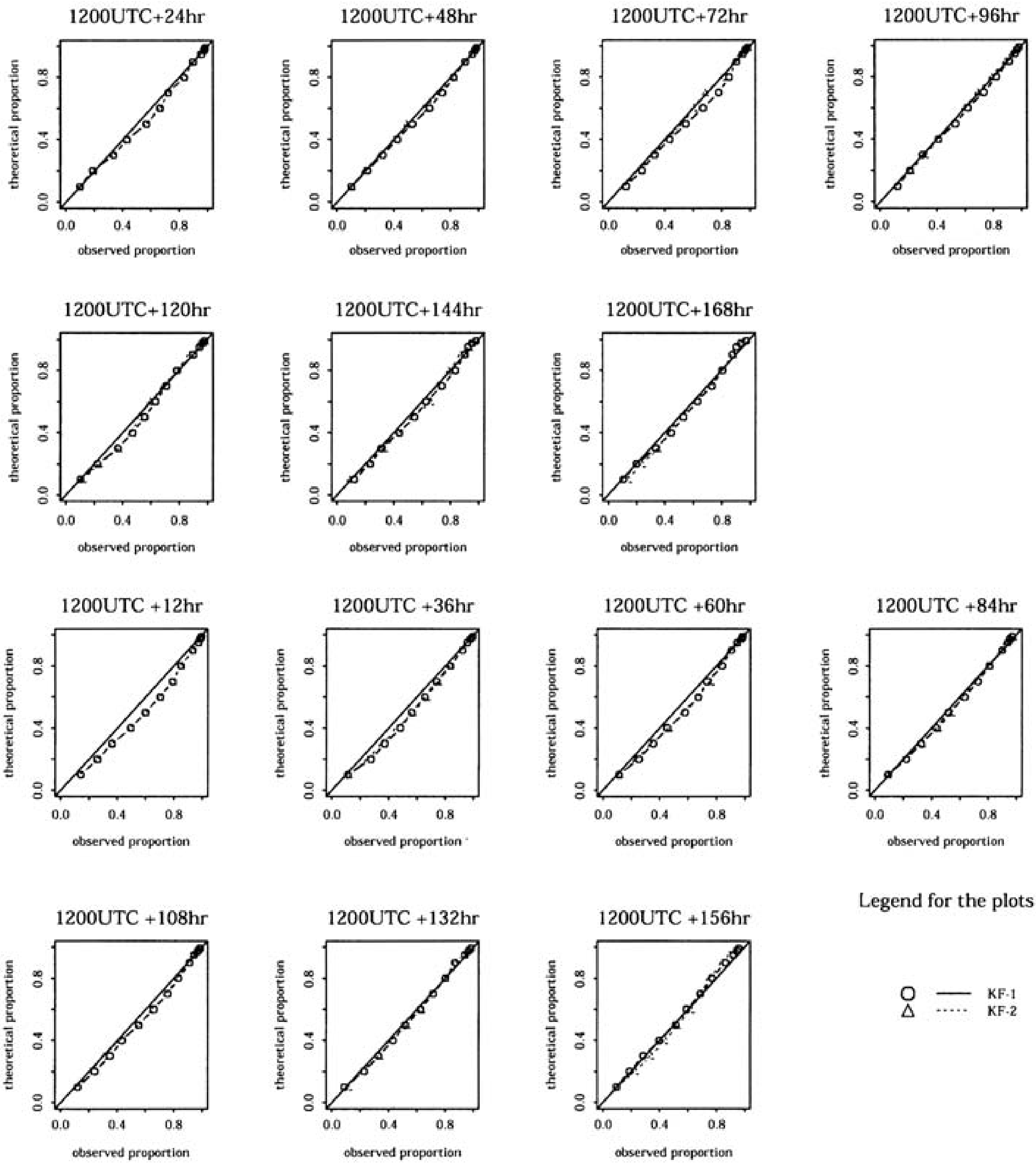

Figure 8. *Average reliability of the prediction intervals for the 2-metre temperature in Akurnes, calculated over the period 1 March 2000–1 March 2001, and presented for each forecast range. The 1:1 line corresponds to perfect reliability. The points falling above the 1:1 line correspond to an underestimation of the size of the theoretical prediction interval. The points falling under the 1:1 line correspond to an overestimation of the size of the theoretical prediction interval.*

is assumed over the whole forecasting period. The KF algorithm as described in Appendix 1 is applied to the shortest forecast ranges valid at different verification times. For the other forecast ranges, the state vector $\hat{X}_{t+k\Delta t|t}$ is not computed but replaced by the one computed for the shortest forecast range valid at the same verification time. Once the state vector $\hat{X}_{t+k\Delta t|t}$ is given, the other parameters are recursively estimated by the KF algorithm as described in Appendix 1, with an updating time lag of $k\Delta t$ or $(k+1)\Delta t$ depending if the forecast is valid at 1200 or 0000 UTC. In practice, the KF algorithm is applied individually to the $+12$ hr and $+24$ hr forecasts and updated every 24 hr. Then, for forecast ranges $k\Delta t > 24$ hr, the state vector $\hat{X}_{t+k\Delta t|t}$ is not computed but replaced by $\hat{X}_{t+12hr|t}$ or by $\hat{X}_{t+24hr|t}$ depending if the forecast is valid at 0000 or 1200 UTC, and then the other parameters are recursively estimated by the adaptive KF algorithm.

3.3. Parameters specification

The parameters are specified for each location as follows:

Y_t is a scalar that defines the observation of T_{2m} or FF_{10} made at time t.

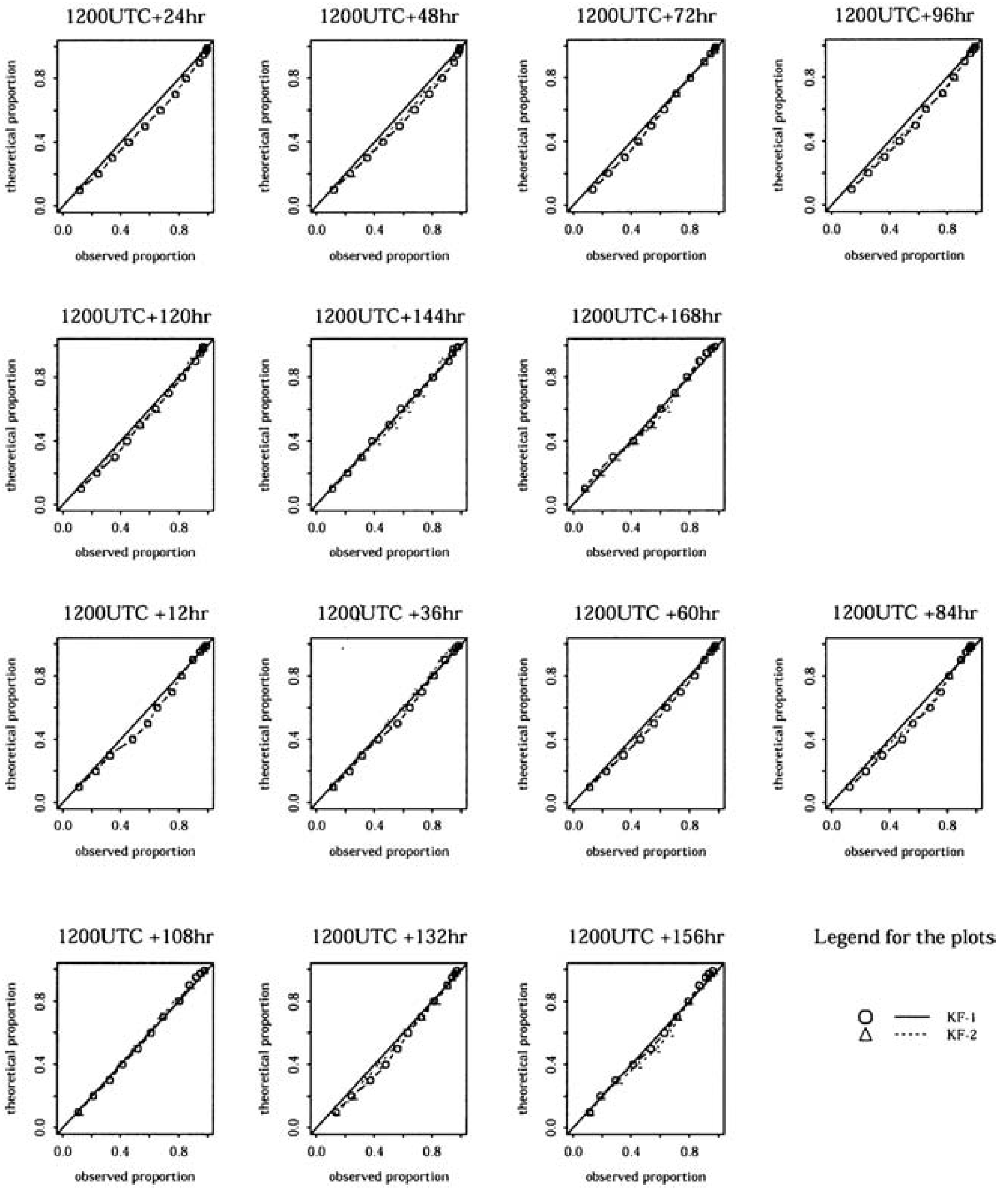

Figure 9. *As Figure 8 but for Stykkisholmur.*

H_t is the 1×2 vector of predictors: $(F_{t|t-k\Delta t},1)$, where $F_{t|t-k\Delta t}$ is the DMO forecast of T_{2m} or FF_{10} made at time $t - k\Delta t$ and valid at time t.
$\hat{X}_{t|t-k\Delta t}$ is a 2×1 vector.
$\hat{X}_{t|t}$ is a 2×1 vector.
Φ is a 2×2 matrix set to the identity matrix I.
Γ is a 2×2 transition matrix set to the identity matrix I.

The initial values for α_t, v_t, V_t, $\hat{X}_{t|t}$ and $P_{t|t}$ (see definition in Appendix 1 and 2) are specified for both T_{2m} and FF_{10} filters as follows:

$$\alpha_0 = 1,\ v_0 = 0,\ V_0 = 1,\ \hat{X}_0 = (1,0),\ P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

3.4. Statistical scores

The performances of the predictions are assessed using the following statistical scores:

$$\text{Mean error:}\quad ME = E[(\varepsilon_{t|t-k\Delta t})] \tag{4}$$

$$\text{Root mean squared error:}\quad RMSE = E[(\varepsilon_{t|t-k\Delta t})^2]^{1/2} \tag{5}$$

where

$$\varepsilon_{t|t-k\Delta t} = \hat{Y}_{t|t-k\Delta t} - Y_t \tag{6}$$

denotes the error in predicting Y_t from time $t - k\Delta t$ and $\hat{Y}_{t|t-k\Delta t}$ denotes the prediction of Y_t made at time $t - k\Delta t$:

$\hat{Y}_{t|t-k\Delta t} = F_{t|t-k\Delta t}$ (Local DMO forecast)
$\hat{Y}_{t|t-k\Delta t} = H_t \hat{X}_{t|t-k\Delta t}$ (KF-1)
$\hat{Y}_{t|t-k\Delta t} = H_t \hat{X}_{t-k\Delta t+24hr|t-k\Delta t}$ (KF-2) valid at 1200 UTC
$\hat{Y}_{t|t-k\Delta t} = H_t \hat{X}_{t-k\Delta t+12hr|t-k\Delta t}$ (KF-2) valid at 0000 UTC

The RMSE of the four different predictions will also be compared to a climatic prediction defined here by the

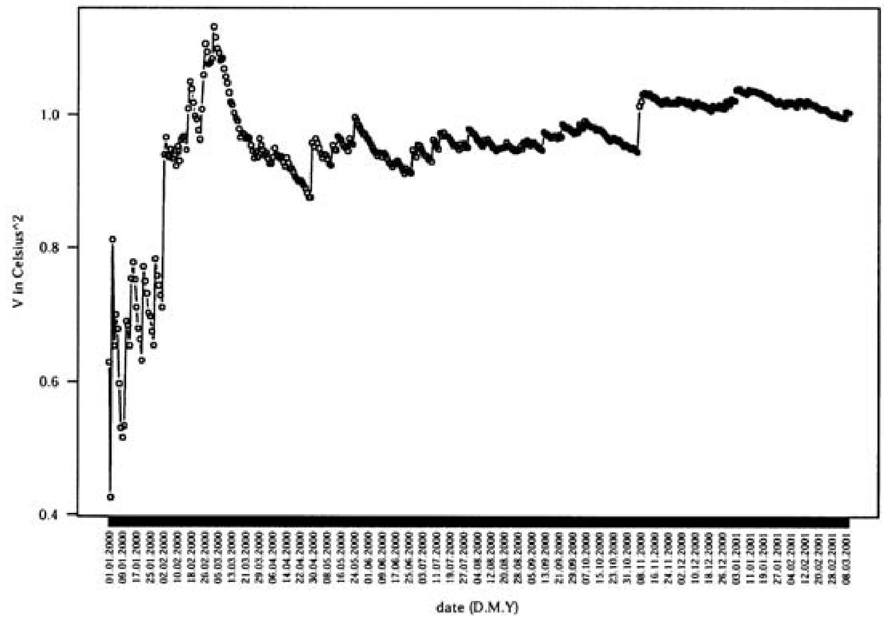

Figure 10. *Two-metre temperature forecast for Reykjavik: Evolution of the observation noise variance V_t, estimated by the algorithm of Smith in the KF-1 protocol and a forecast range of 24 hr.*

30-year mean daily T_{2m} or FF_{10} value over the period 1971–2000. The reliability of the prediction interval (3) is assessed by counting the proportion of times the observations belong to the prediction interval, over the verification period. If the $100(1-\alpha)\%$ prediction interval is reliable on average, the proportion of observations Y_t lying within it should be $100(1-\alpha)\%$. If the prediction interval is too large, the proportion of observations Y_t lying within it should be greater than $100(1-\alpha)\%$. If the prediction interval is too small, the proportion of observations Y_t lying within it should be lower than $100(1-\alpha)\%$. For FF_{10}, if the lower limit of the prediction interval is negative, the limit is set to 0.

3.5. Preliminary results

In studying the methodology, it was observed that in the presence of large forecasting errors (mainly for ranges larger than 96 hr), the Jazwinski algorithm could produce high values for β_t (Eq. C-3, Appendix 3) and consequently make the filter quite sensitive and produce oscillations in the prediction (alternately positive and negative errors with a large magnitude) and high predictive variances. As a consequence, the performances of the KF-1 protocol were observed to be unbiased but relatively poor in terms of RMSE for forecast ranges larger than 96 hr. The smoothing procedure proposed by Jazwinski was not considered here, but in order to limit this instability problem, it was found effective simply to define an upper limit for β_t chosen with a certain degree of arbitrariness among a small set of values in order to lower the RMSE to a minimum value. The use of an upper limit for β_t is described as follows. Instead of opening the filter wide in reaction to a large error, the adaptation is spread over several time steps if necessary. If the large error is an isolated case or is due perhaps to an observational error, the influence of this event is relatively limited. If the error is persistent, the filter will progressively open to this change in a controlled manner. The KF-1 protocol was observed to be more sensitive to the choice of this upper limit than the KF-2 protocol, especially when filtering FF_{10}. This is due to the fact that both $\hat{X}_{t+k\Delta t|t}$ and $P_{t+k\Delta t|t}$ are dependent upon the choice of this upper limit. In the KF-2 protocol, on the other hand, $\hat{X}_{t+k\Delta t|t}$ is not computed directly for forecast ranges greater than 24 hr, but given by the KF corresponding to the two shortest forecast ranges valid at 0000 or 1200 UTC. This fixed upper limit was observed to be reached more frequently as the forecast projection time increased, between 5% and 30% of the time. The following results present the performances of the adaptive KF with an upper limit for β_t set to 0.2 for both T_{2m} and FF_{10}.

3.6. Examples of the behaviour of the adaptive KF

Figures 2 and 3 present respectively the predictions of T_{2m} for Akureyri and Stykkisholmur issued on 15 March 2000 for the following seven days, together with the 80% prediction intervals (PI). Figures 4 and 5 give the predictions of FF_{10} for Hveravellir and Reykjavik issued the same day. The observed values are also represented for comparison. For Akureyri (Figure 2), the T_{2m} DMO forecasts present a marked negative bias. The KF has noticed this systematic bias and attempts to correct it. The result is a prediction systematically warmer than the DMO one. The KF predictions are more accurate than the DMO predictions 11 times out of 14. The size of the 80%

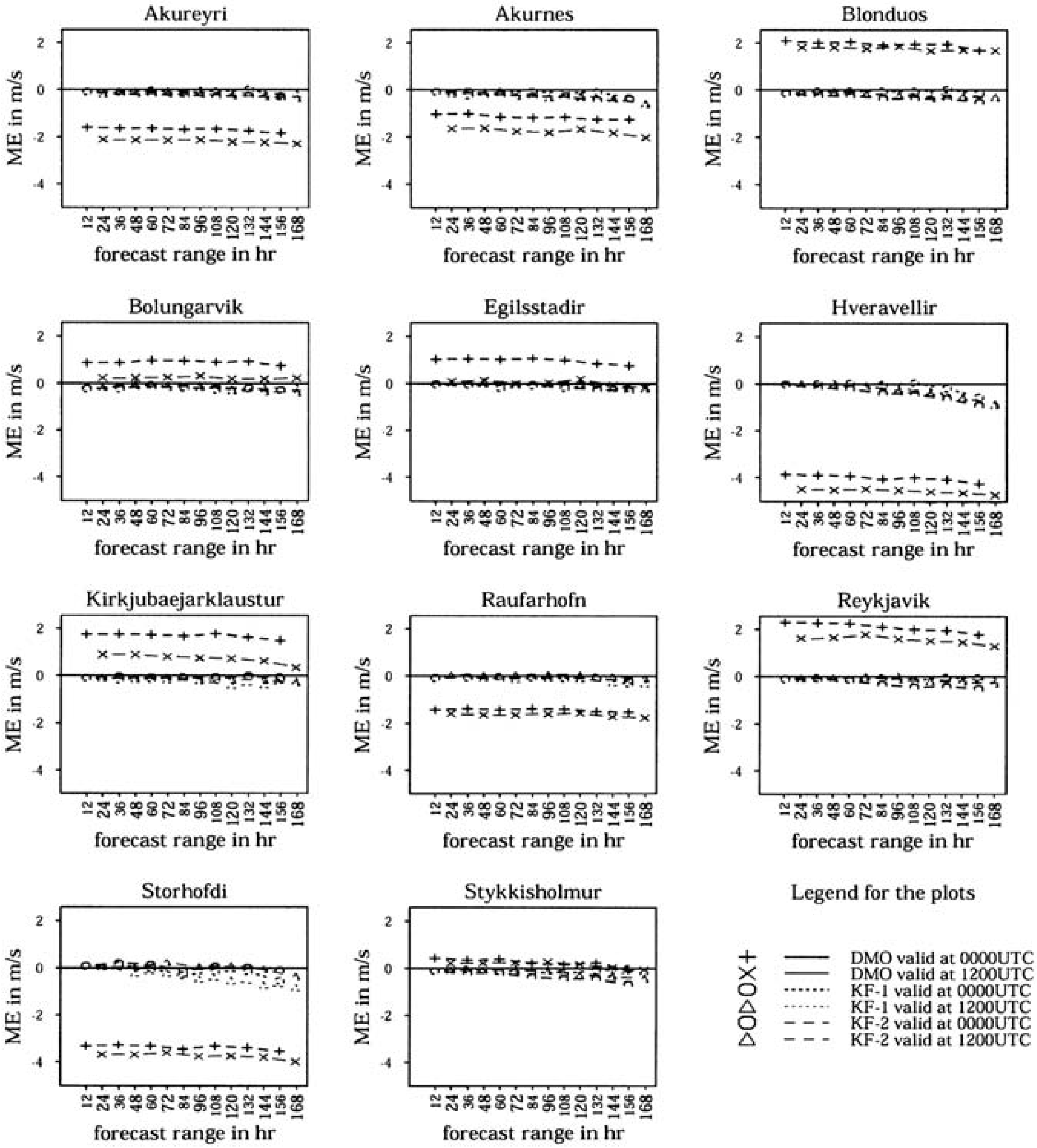

Figure 11. *Mean error (ME) for the 10-metre wind-speed forecasts, calculated for each location over the period 1 March 2000–1 March 2001 and presented as a function of the forecast range.*

PI increases slowly with the forecast range, showing an increase of uncertainty. The 80% uncertainty in this case represents about +/− 3 °C up to + 96 hr range and then slowly increases up to +/− 7 °C Celsius at + 168 hr range. It can also be seen that all the observations except three belong to the 80% PI, i.e. 78.6% of the data. For Stykkisholmur (Figure 3), the DMO forecast does not exhibit any systematic bias. The KF algorithm has noticed that and produces a prediction that oscillates around the DMO value. In such a case, the KF prediction does not perform better than the DMO one. The size of the 80% PI is quite narrow, about +/− 2 °C up to + 96 hr range, and then increases slowly, up to +/− 6 °C at + 168 hr range. There are 11 observations out of 14 belonging to the 80% PI, i.e. 78.6% of the data. For Hveravellir (Figure 4), the FF_{10} DMO forecast has the tendency to underestimate the observed wind speed. The KF predictions are systematically increased. This systematic adjustment turns out to be successful up to + 96 hr range and produces a more accurate prediction than the DMO one. Then the observed wind speed drops down and the DMO forecast becomes better than the KF prediction. In total, over the seven forecasting days, the KF predictions are closer to the observations nine times out of 14. The size of the 80% PI is very large in this case, more than +/− 8 m/s showing the difficulty in predicting such an outcome. It is observed however that ten observations are within the 80% PI limits, i.e. 71% of the data. For Reykjavik (Figure 5), the DMO is overestimating FF_{10}, most of the time. The KF reduces this bias and the adjustment produces a more accurate prediction ten times out of 14. All the observations except three are within the limit of the 80% PI, i.e. 78.6% of the data.

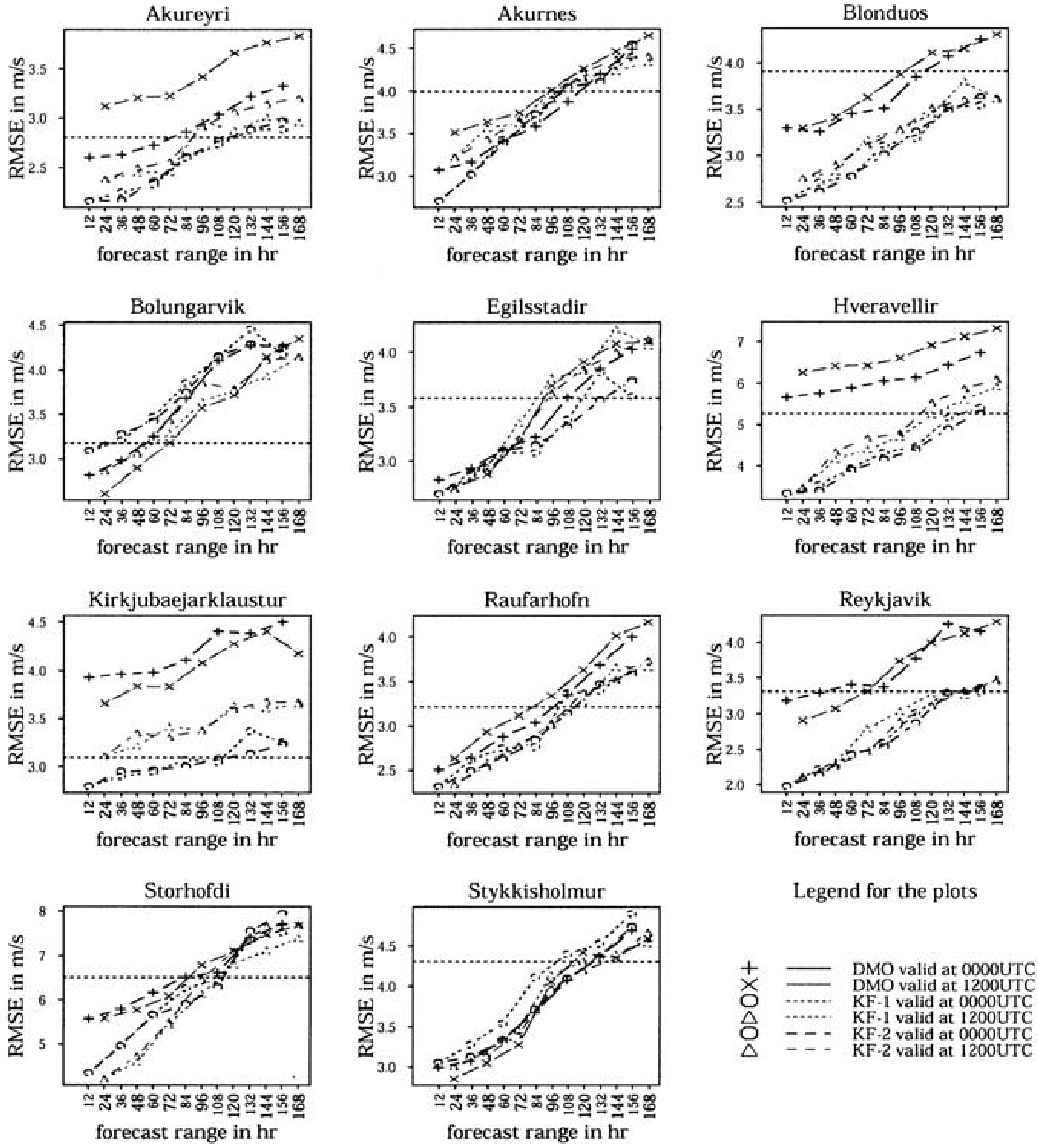

Figure 12. *Root mean squared error (RMSE) for the 10-metre wind-speed forecasts, calculated for each location over the period 1 March 2000–1 March 2001 and presented as a function of the forecast range. The horizontal dashed line gives the RMSE for a climatic prediction.*

3.7. Summary statistics

In this section, the quality of the adaptive KF is studied over a one-year period, from 1 March 2000 to 1 March 2001.

3.7.1. 2-metre temperature

Figures 6 and 7 present for each location the evolution of ME and RMSE respectively, as a function of the forecast range. The two sets of KF predictions remove the bias (ME), if any, at all forecast ranges, but do not necessarily reduce the RMSE. This shows that the KF predictions are unbiased but not always more accurate than the DMO forecasts. The reduction of RMSE is observed at locations where the bias is quite systematic. For locations where the error is not systematic and more random, the RMSE from the KF predictions are similar to the DMO ones. The KF-2 protocol quite often gives slightly lower RMSE than the KF-1 protocol for forecast ranges larger than 24 hr. A comparison with a climatic prediction shows the superiority of the NWP model up to a range of seven days for the sites where the DMO is unbiased. For locations where the DMO forecast displays a strong systematic bias, the climatic prediction displays a lower RMSE than the DMO prediction beyond the 96 hr-range. The KF prediction is unbiased at all sites and displays a RMSE lower than the climatic prediction at all forecast ranges, except for Blonduos beyond 96 hr. The reliability of the prediction intervals is observed to be remarkable for both protocols, at all forecast ranges and

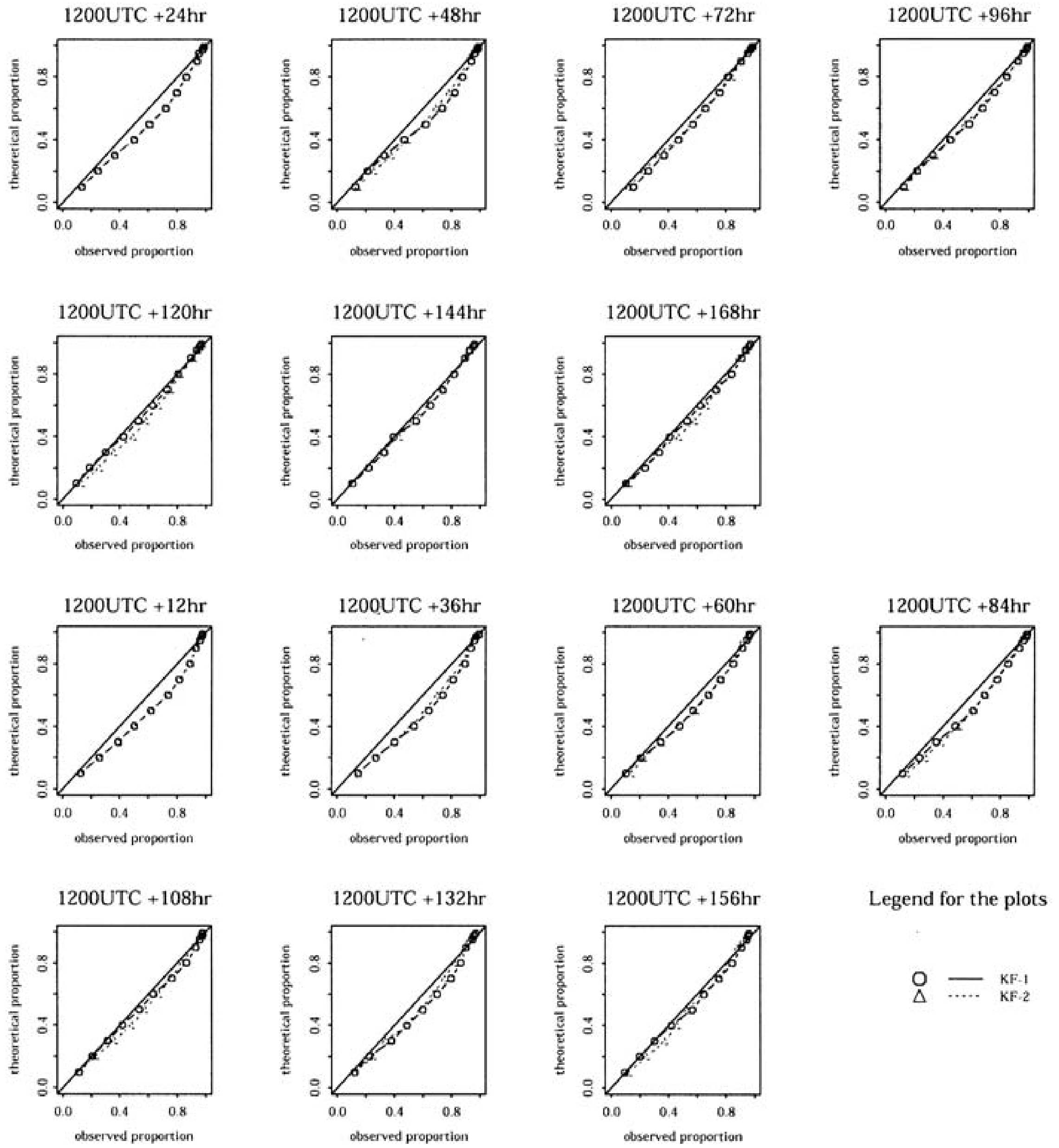

Figure 13. *Average reliability of the prediction intervals for the ten-metre wind-speed in Reykjavik, calculated over the period 1 March 2000–1 March 2001, and presented for each forecast range. The 1:1 line corresponds to perfect reliability. The points falling above the 1:1 line correspond to an underestimation of the size of the theoretical prediction interval. The points falling under the 1:1 line correspond to an overestimation of the size of the theoretical prediction interval.*

all locations. There is, in some cases, a tendency towards a slight overestimation of the size of the prediction intervals. In such cases, the proportion of observations belonging to the prediction interval is slightly larger than expected. An example of this reliability is given for Akurnes in Figure 8 and Stykkisholmur in Figure 9. Figure 10 illustrates an example of the evolution of V_t for Reykjavik estimated by the Smith algorithm (see Appendix 2). During the first month, the series fluctuates and presents jumps. After that the evolution of V_t is smooth and the series reaches a rather stationary level. It is worth mentioning that it is observed (but not shown) that V_t increases with the forecast range as a consequence of the increase of prediction uncertainty.

3.7.2. 10-metre wind-speed

Figures 11 and 12 present the evolution of ME and RMSE respectively, as a function of the forecast range. The two sets of KF predictions are equally efficient in removing any bias (ME). Concerning the accuracy of the forecasts, the results are station dependent. For the stations where the error is quite systematic, both KF protocols reduce the RMSE but the advantage of using one protocol rather than the other is not clearly defined. For the stations where the error is more random, the KF predictions exhibit an accuracy similar to the DMO forecasts, except for Bolungarvik and Stykkisholmur where it is slightly larger. Depending on the station,

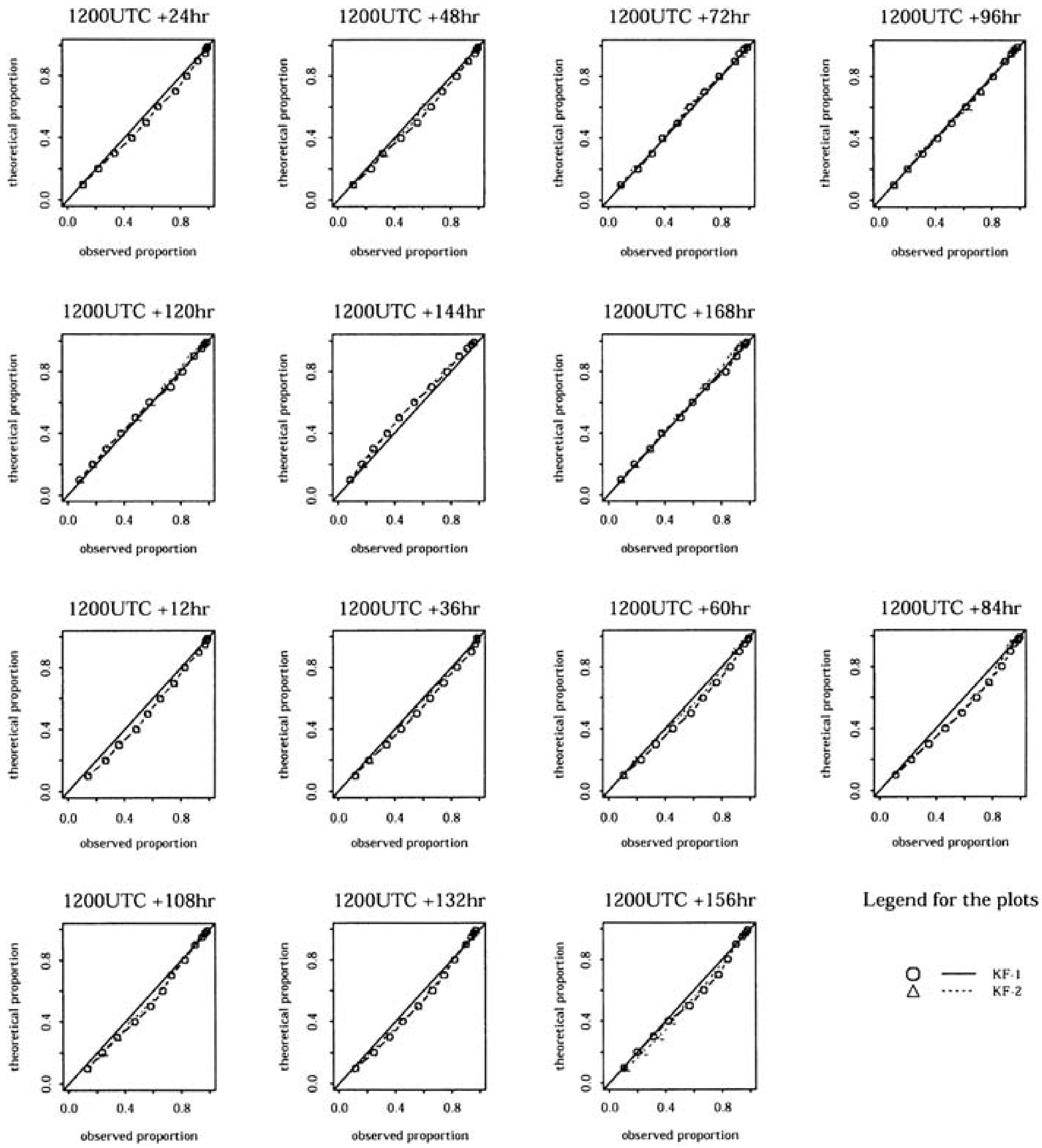

Figure 14. *As Figure 13 but for Egilsstadir.*

the climatic prediction displays a better accuracy than the NWP beyond ranges exceeding 3–5 days. For Kirkjubaejarklaustur, at 1200 UTC verification time, the climatic prediction gives lower RMSE than both DMO and KF predictions at all forecast ranges. These results also show the difficulty of verifying and Kalman filtering forecasts at locations where the measurements are quite uncertain because they are made visually using the Beaufort scale. Here too the reliability of the prediction intervals is observed to be quite remarkable for both protocols, at all forecast ranges and all the locations. In some cases there is a tendency towards a slight overestimation of the size of the prediction intervals. An example is presented for Reykjavik in Figure 13, and Egilsstadir in Figure 14. Figure 15 gives an example of the evolution of V_t for Akureyri. The fluctuations and the jumps are quite frequent during the first month. After that the evolution becomes quite smooth and the series reaches a rather steady level. Here too, as for T_{2m}, it is observed (but not shown) that V_t increases with the forecast range.

3.8. Discussion

In weather forecasting, the Kalman filter is commonly used as a tool to adjust the local forecasts of surface parameters such as T_{2m} and FF_{10} in order to remove the systematic bias. Very little use is made of the predictive variance (Eq. A-5 Appendix 1) and the assessment of the performance of such a method is usually limited to the study of the bias and the accuracy of the prediction, but not the reliability. It is possible to manually adjust the sensitivity of the KF by tuning the observation and system noise variances. However, this should be done in relation to the forecast range. A high sensitivity may be advantageous for short ranges

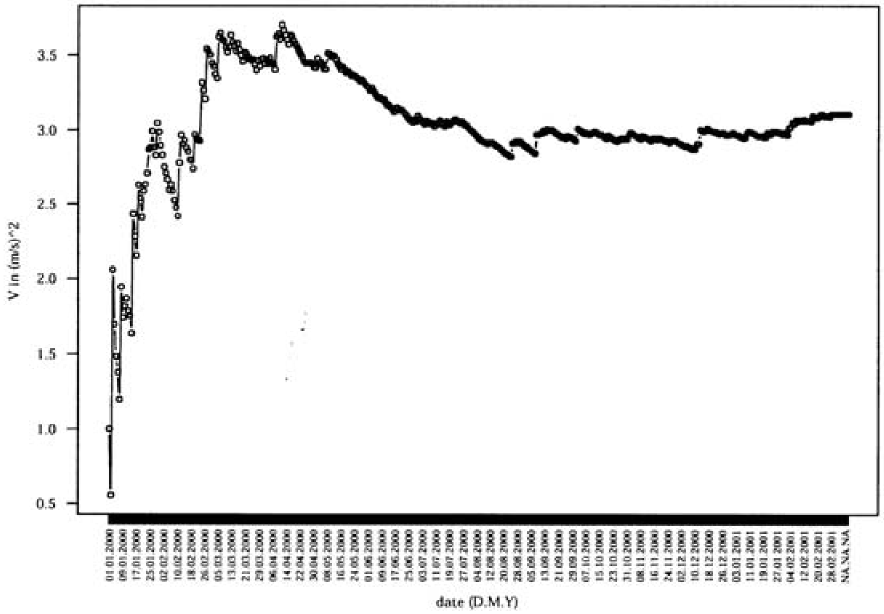

Figure 15. *Ten-metre wind-speed forecast in Akureyri: Evolution of the observation noise variance V_t, estimated by the algorithm of Smith in the KF-1 protocol and a forecast range of 24 hr.*

but is not necessarily a good feature for longer ranges because of the delay in the response, especially when the errors are more random than systematic. Finding the right balance between adaptivity, reliability and forecast range may require tedious efforts when many sites are considered. The KF procedure presented here was developed with the following goals: reduction of the systematic errors, automatic adaptation to abrupt and/or structural changes in relation to the forecast range, and reliability of the prediction intervals. From the information given by the predictive distributions, the forecaster may then issue the appropriate forecast to a given situation. Even in situations where the KF predictions are observed not to be more accurate than the DMO forecasts, the adaptive KF procedure still conveys useful information about the uncertainty of the local predictions. Both protocols presented here were observed to be unbiased and reliable on average. It was also observed how difficult it is to improve the accuracy of the forecasts in situations where the DMO is already unbiased. It is difficult to clearly favour one protocol rather than another because the results are not fundamentally different. The results also demonstrated the superiority of a Kalman filtered NWP forecast over a climatic prediction, up to 7 days ahead for T_{2m}, and 3 to 5 days ahead for FF_{10}.

4. Summary and conclusions

The potential of a statistical method such as the KF is not limited to bias reduction of a local forecast. It can also provide guidance on the prediction uncertainty in the form of an interval, reliable in average at a particular location. The design of a reliable KF procedure can prove to be difficult without any prior knowledge of the noise statistics. An adaptive KF procedure is an attractive solution to this because it can be set up quite rapidly and applied individually at as many locations as required. Further development might incorporate more sophisticated modelling of the different matrices involved in the system equation, set here to the identity matrix. The algorithms used in this study to adaptively estimate the noise statistics are not unique and other solutions described in the literature could also be considered.

Acknowledgements

The author wishes to thank Trausti Jónsson, Sigrún Karlsdóttir, Guðmundur Hafsteinsson and the reviewer for helpful comments and suggestions on the draft of this paper.

References

Bolzern, P., Ferrario, M. & Fronza, G. (1980) Adaptive real-time forecast of river flow-rates from rainfall data. *J. Hydrology* **47**: 251–267.

De Meyer, F. (2000) Optimal linear parameter estimation in dynamic stochastic models. *Publication scientifique et technique* no. 10, Institut Royal Météorologique de Belgique, 217 pp.

Golanis, G. & Anadranistokis, M. (2002) A one-dimensional Kalman filter for the correction of near surface temperature forecasts. *Meteorol. Appl.* **9**: 437–441.

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

Harrison, P. J. & Stevens, C. F. (1976) Bayesian forecasting. *J. R. Stat. Soc.* (Ser. B) **38** (3): 205–247.

Homleid, M. (1995) Diurnal corrections of short-term surface temperature forecasts using the Kalman filter. *Wea. Forecasting* **10** (4): 689–707.

Jazwinski, A. H. (1969) Adaptive filtering. *Automatica* **5**: 475–485.

Kalman, R. E. (1960) A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82**: 34–45.

Klein, W. H. & Lewis, F. (1970) Computer forecasts of maximum and minimum temperatures. *J. Appl. Meteorol.* **9**: 350–359.

Mehra, R. K. (1970) On the identification of variances and adaptive Kalman filtering. *IEEE Transactions on Automatic Control*, AC-15, no. 2, 175–184.

Meinhold, R. J. & Singpurwalla, N. D. (1983) Understanding the Kalman filter. *Am. Statistician* **37** (2): 123–127.

Persson, A. (1991). 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 Phenomena and Their Verification, Wageningen, the Netherlands, WMO PSMP Report Series no. 34, XX-27-XX-32.

Sage, A. P. & Husa, G. W. (1969) Algorithms for sequential adaptive estimation of prior statistics. *Proc. 8th IEEE Symposium on Adaptive Processes*, Pennsylvania State University, University Park, PA.

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 PSMP Report Series no. 34, XX-33-XX-37.

Smith, G. L. (1967) Sequential estimation of observation error variances in a trajectory estimation problem. *AIAA J.* **5** (11): 1964–1970.

West, M., Harrison, P. J. & Migon, H. S. (1985) Dynamic generalized linear models and bayesian forecasting. *J. Am. Stat. Assoc.* **80** (389): 73–97.

Wilson, L. J. & M. Vallée (2002) The Canadian Updatable Model Output Statistics (UMOS) system: design and development tests. *Wea. Forecasting* **17** (2): 206–222.

Wilson, L. J. & M. Vallée (2003) The Canadian Updatable Model Output Statistics (UMOS) system: validation against Perfect Prog. *Wea. Forecasting* **18** (2): 288–302.

WMO (1988) *Manual on Codes. Vol. 1. International Codes.* WMO no 306.

Appendix 1: The Recursive Kalman filter algorithm

For a given time lag $k\Delta t$ ($k > 0$), the recursive algorithm estimates the following parameters:

$\hat{X}_{t|t-k\Delta t}$ is the prior estimate of the state vector X_t given information up to time $t-k\Delta t$:

$$\hat{X}_{t|t-k\Delta t} = \Phi \hat{X}_{t-k\Delta t|t-k\Delta t} \quad \text{(A-1)}$$

Where Φ is a known $n \times n$ transition matrix.

$P_{t|t-k\Delta t}$ is the prior error covariance matrix of X_t ($Var(X_t - \hat{X}_{t|t-k\Delta t})$) given information up to time $t-k\Delta t$:

$$P_{t|t-k\Delta t} = \Phi P_{t-k\Delta t|t-k\Delta t} \Phi^T + \Gamma W_t \Gamma^T \quad \text{(A-2)}$$

Where Γ is a $n \times r$ matrix of fixed coefficients and W_t the system noise variance.

Using all the available information up to and including time $t - k\Delta t$, the mean $\hat{\mu}_{t|t-k\Delta t}$ and variance $\hat{\sigma}^2_{t|t-k\Delta t}$ of the predictive distribution of any observation Y_t, are derived as follows (see for instance Harrison & Stevens 1976):

$$(Y_t | \hat{X}_{t|t-k\Delta t}, H_t) \sim N\left(\hat{\mu}_{t|t-k\Delta t}, \hat{\sigma}^2_{t|t-k\Delta t}\right) \quad \text{(A-3)}$$

where H_t represents the predictors or explanatory variables.

$$\hat{\mu}_{t|t-k\Delta t} = H_t \hat{X}_{t|t-k\Delta t} \quad \text{(A-4)}$$

and

$$\hat{\sigma}^2_{t|t-k\Delta t} = H_t P_{t|t-k\Delta t} H_t^T + V_t \quad \text{(A-5)}$$

where V_t is the observation noise variance.

Once the observation Y_t is available, the KF algorithm updates the prior estimates. $\hat{X}_{t|t}$ is the posterior estimate of X_t given information up to time t:

$$\hat{X}_{t|t} = \hat{X}_{t|t-k\Delta t} + K_t e_{t|t-k\Delta t} \quad \text{(A-6)}$$

where

$$e_{t|t-k\Delta t} = Y_t - \hat{\mu}_{t|t-k\Delta t} \quad \text{(A-7)}$$

is the error in predicting Y_t using information up to and including time $t-k\Delta t$, and

$$K_t = P_{t|t-k\Delta t} H_t^T \left(H_t P_{t|t-k\Delta t} H_t^T + V_t\right)^{-1} \quad \text{(A-8)}$$

is the Kalman gain matrix.

$P_{t|t}$ is the posterior error covariance matrix of X_t (Var $(X_t - \hat{X}_{t|t})$) given information up to time t:

$$P_{t|t} = (I - K_t H_t) P_{t|t-k\Delta t} \quad \text{(A-9)}$$

where I denotes the identity matrix.

Appendix 2: Estimation of the observation noise variance V_t, using Smith's method

The algorithm of Smith (1967) allows the adaptive identification of the observation noise variance V_t at each time step. In this method, V_t is assumed to be the product of a nominal value V_0 and a coefficient α_t, which is assumed to have an inverted gamma distribution with parameters δ_t and υ_t. This method is quite similar to the learning procedure described in West et al. (1985). After making the approximation $\delta_t = \alpha_t$, the adaptive estimation of V_t is obtained as follows:

$$V_t = \hat{\alpha}_{t|t-k\Delta t} V_0 \quad \text{(B-1)}$$

$$\hat{\alpha}_t = \frac{\hat{\alpha}_{t|t-k\Delta t}}{v_{t|t-k\Delta t}+1}\left\{ v_{t|t-k\Delta t} + \frac{e^2_{t|t-k\Delta t}}{H_t P_{t|t-k\Delta t} H_t^T + V_t} \right\} \quad \text{(B-2)}$$

$$v_t = v_{t|t-k\Delta t} + 1 \quad \text{(B-3)}$$

Where $\hat{\alpha}_{t|t-k\Delta t}$ and $v_{t|t-k\Delta t}$ are the prior estimates of α_t and v_t respectively, at time $t-k\Delta t$.

Appendix 3: Estimation of the system noise variance W_t, using Jazwinski's method

The method of Jazwinski (1969) adjusts the system noise variance W_t in order to produce consistency between the residuals and their statistics and to avoid the KF becoming divergent. This method was originally used in orbit determination and found later applications in other sciences, such as hydrology (e.g. for adjusting river flow-rates forecasts; see Bolzern et al. 1980). A description of the Jazwinski algorithm can also be found in De Meyer (2000). The method defines

$$W_t = \beta_t I \quad \text{(C-1)}$$

where I is the identity matrix and β_t the coefficient to be determined according to the following requirement:

$$e^2_{t|t-k\Delta t} < E\left[e^2_{t|t-k\Delta t} | W = 0\right]$$

where

$$E\left[e^2_{t|t-k\Delta t} | W = 0\right] = H_t \Phi P_{t-k\Delta t|t-k\Delta t} \Phi^T H_t^T + V_t \quad \text{(C-2)}$$

This leads to the following estimator

$$\beta_t = \begin{cases} \dfrac{e^2_{t|t-k\Delta t} - E\left[e^2_{t|t-k\Delta t} | W = 0\right]}{H_t \Gamma \Gamma^T H_t^T} & \text{if positive} \\ 0 & \text{otherwise} \end{cases} \quad \text{(C-3)}$$

This estimate is based on one residual. The statistical significance of β_t can be improved, if necessary, by using a smoothing procedure consisting of averaging the k previous estimates. If a change has occurred and a divergence is observed in the predictions, an action is taken in order to increase the prior covariance matrix $P_{t|t-k\Delta t}$ and consequently to allow the model parameter X_t to adapt faster. With the Jazwinski procedure, $P_{t|t-k\Delta t}$ is estimated once the observation becomes available at time t. However, the goal of this paper is to estimate the parameters of the predictive distribution (both the mean and the variance, Eq. A-4 and A-5 Appendix 1) before the observation is made, and to produce at time $t-k\Delta t$ a prediction interval valid at time t. From this requirement, the prior error covariance matrix $P_{t|t-k\Delta t}$ needs to be estimated at time $t-k\Delta t$ and also W_t. In order to provide this information, the Jazwinski algorithm is used in this study to arbitrarily define W_t a priori as follows:

$$W_t = \beta_{t-k\Delta t} I \quad \text{(C-4)}$$

The system noise variance W_t, to be used in the next prediction valid at time t, is defined here a priori at time $t-k\Delta t$ according to the magnitude of the error made in the previous forecast. In doing so, the prior error covariance matrix $P_{t|t-k\Delta t}$ is inflated, and consequently the predictive variance $\hat{\sigma}^2_{t|t-k\Delta t}$ as well. It indicates that a large error has been detected and the sensitivity of the Kalman filter is increased for the next adjustment.