Proceedings of the Second IASTED International Multi-Conference
SIGNAL AND IMAGE PROCESSING (ACIT-SIP)
June 20-24, 2005, Novosibirsk, Russia

IMPROVING LOCAL WEATHER FORECASTS FOR AGRICULTURAL APPLICATIONS

T.G. Doeswijk
Systems and Control Group
Wageningen University
P.O. Box 17
6700 AA Wageningen, THE NETHERLANDS
timo.doeswijk@wur.nl

K.J. Keesman
Systems and Control Group
Wageningen University
P.O. Box 17
6700 AA Wageningen, THE NETHERLANDS
karel.keesman@wur.nl

ABSTRACT

For controlling agricultural systems, weather forecasts can be of substantial importance. Studies have shown that forecast errors can be reduced in terms of bias and standard deviation using forecasts and meteorological measurements from one specific meteorological station. For agricultural systems usually the forecasts of the nearest meteorological station are used whereas measurements are taken from the systems location. The objective of this study is to evaluate the reduction of the forecast error for a specific agricultural system. Three weather variables , that are most relevant for greenhouse systems are studied: temperature, wind speed, and global radiation. Two procedures are used consecutively: diurnal bias correction and local adaptive forecasting. For each of the variables both bias and standard deviation were reduced. In general, if local measurements are reliable, forecast errors can be reduced considerably.

KEY WORDS

weather forecast, Kalman filter, forecast error, bias, standard deviation

1 Introduction

In many agricultural systems all kinds of weather variables, such as temperature, radiation and rain, have a dominant effect on the systems' behavior. Weather input variables are not only a disturbance to the system but are also a resource, (*e.g.* global radiation drives plant growth). For maximization of plant production some of the control inputs should closely follow changes in weather conditions, *e.g.* CO_2-dosing in green-houses should anticipate on changes of global radiation. Therefore, when controlling agricultural systems, weather forecasts can be of substantial importance, especially when anticipating control strategies are used. [see *e.g.* 1, 2]. If forecasts of less than one hour are used, the so-called "lazy man" weather prediction, where the forecast is chosen equal to the most recently measured value, seems to be reasonable [3]. If, however, the forecast horizon increases, preferably commercial forecasts should be used.

It is well known that, because of the chaotic behavior of the atmosphere, weather forecasts can be rather uncertain. This uncertainty increases as the forecast horizon increases. Many efforts are taken by meteorologists to improve the quality of the weather forecasts, but forecasts of weather variables remain uncertain. For instance, the 2 meter temperature forecast has a variance of 2 $^oC^2$ for the zero-hour ahead forecast in "De Bilt", The Netherlands [4].

Previous research has been done to improve local weather forecasts. For instance, biases can be largely removed from meteorological model outputs [5]. This bias reduction procedure uses a Kalman filter that predicts a diurnal forecast error. This forecast error then has to be added to the original forecast. The procedure is further called "diurnal bias correction" (DBC). In specific cases like mountainous areas and places surrounded by seas [6] this DBC proved to reduce the bias drastically. Standard deviations of the forecast error however, are not reduced with this procedure. However, reduction of the standard deviation of the forecast error can be obtained as well [4]. The standard deviation reduction procedure uses local measurements in a Kalman filter to update the forecast. This procedure is further referred to as "local adaptive forecasting" (LAF). The best performance was obtained obtained by applying both procedures: first bias reduction, then standard deviation reduction [4].

The presented studies in [5] and [4] used meteorological data of specific meteorological stations. Measurements at these meteorological stations need to fulfill specific meteorological requirements [7]. In general, local measurements for agricultural systems do not fulfill the meteorological requirements, but they do represent the actual local situation. Furthermore, the forecasts from weather agencies are specific for some predetermined places, but most often the specific agricultural systems are not located at these places. At last, local circumstances such as soil properties, different altitudes, presence of water resources (lake, sea) etc., have an effect on local weather conditions and thus should be accounted for by the local weather forecast. Therefore it is assumed that weather forecasts for agricultural systems can be improved for both bias and standard deviation.

The purpose of this paper is to show that local forecasts for agricultural systems can be improved by using local measurements. This improvement is based on reduction of the *bias* as well as the *standard deviation* of the forecast error.

In section 2 theoretical background is given and the analyzed data are explained. In section 3 the results of three weather variables are presented: temperature, wind speed and global radiation. The results are discussed in section 4. Finally, some conclusions are presented in section 5.

2 Background

2.1 Kalman filtering

Both procedures, DBC and LAF, are based on a discrete-time state-space system representation of either forecast errors or forecasts and use the Kalman filter as the main algorithm to update weather forecasts.

The stochastic discrete-time state-space system used in both procedures is given by:

$$x(k+1) = A(k)x(k) + B(k)u(k) + G(k)w(k) \quad (1)$$
$$y(k) = C(k)x(k) + v(k) \quad (2)$$

where $x(k) \in \mathbb{R}^{M+1}, u(k) \in \mathbb{R}^{M+1}$, with M the maximum forecast horizon, and $y(k) \in \mathbb{R}^p$, with p the dimension of the actual output vector. It is assumed that the disturbance input $w(k)$ (so called "system noise") and measurement noise $v(k)$ are zero-mean Gaussian random sequences with:

$$E\big[w(k)\big] = 0, \qquad E\big[w(k)w^T(k)\big] = Q \quad (3)$$
$$E\big[v(k)\big] = 0, \qquad E\big[v(k)v^T(k)\big] = R \quad (4)$$

The matrices $A(k), B(k), C(k), G(k), Q$ and R are system dependent and will be defined in the subsequent sections.

As mentioned before, the so-called Kalman filter is used to update local weather forecasts. For an elaborate description of the Kalman filter we refer to Gelb [8]. Here, the algorithm is briefly outlined. Given a system in state-space form (1)-(2) with noise properties (3)-(4) and a measured output $y(k)$, the Kalman filter estimates the states at time instance k with the smallest possible error covariance matrix. The following Kalman filter equations for the discrete-time system (1)-(2) are used to estimate the new states (updated forecast errors in DBC or forecasts in LAF) when new observations become available:

$$\hat{x}(k+1|k) = A(k)\hat{x}(k|k) + B(k)u(k) \quad (5)$$
$$P(k+1|k) = A(k)P(k|k)A(k)^T + G(k)QG(k)^T \quad (6)$$
$$K(k+1) = P(k+1|k)C^T \big[CP(k+1|k)C^T + R\big]^{-1} \quad (7)$$
$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1) \big[y(k+1) - C\hat{x}(k+1|k)\big] \quad (8)$$
$$P(k+1|k+1) = P(k+1|k) - K(k+1)CP(k+1|k) \quad (9)$$

where $\hat{x}(k+1|k)$ denotes the estimate of state x at time instant $k+1$ given the state at k, and $A(k)^T$ is the transpose of $A(k)$. Furthermore, $K(k+1)$, known as Kalman gain, denotes the weighting matrix related to the prediction error $\big[y(k+1) - C\hat{x}(k+1|k)\big]$.

2.2 Diurnal bias correction

It has been shown by Homleid [5] that systematic errors from numerical weather prediction models can be largely removed. A brief outline of the algorithm is given.

The basic assumption is that the prediction errors are assumed to vary only in a 24 hour context. The states $x_1 \cdots x_{24}$ represent the forecast errors at times from 0000 UTC until 2300 UTC. No input is present in this system. The output y is defined by the measurement at a specific time and is a scalar. The system matrices are given by: $A = I$, $G = I$ and C is time-varying *e.g.* $C = [1\ 0 \cdots 0]$ at 0000 UTC, $C = [0\ 1\ 0 \cdots 0]$ at 0100 UTC *etc*. The Kalman filter matrices Q and R are time-invariant where Q is a symmetric Toeplitz matrix with ones on the diagonal, premultiplied by the variance W^2 and $R = V^2$ the variance of the measurement error. The W/V-ratio determines the update rate. The optimal W/V-ratio can vary between meteorological stations and can vary between weather variables [e.g. 5, 4]. The initial covariance matrix $P(0)$ is typically chosen as: $P(0) = 10^6 I$, with I the identity matrix. The estimated prediction errors ($\hat{x}(k{+}1|k{+}1)$ are added to the external forecasts independent of the forecast horizon.

2.3 Local adaptive forecasting

Local measurements are used to update the short term forecasts. The updating algorithm as described by Doeswijk and Keesman [4] uses a linear, time-varying system in state-space form that describes the evolution of the forecasts. Every stochastic l-steps ahead forecast at time instant k is treated as a state variable $x_l(k)$ where $l = 0, ..., M$, with M the maximum forecast horizon. Consequently, $x_0(k)$ represents the actual state, $x_1(k)$ the one-step ahead forecast, *etc*. As a consequence, the state vector x always represents the forecast horizon (from 0 to $M - k$ hours ahead). Subsequently, the effective system dimensions reduce as long as there are no new external forecasts available. The external forecasts that become available at time instant k^* are treated as deterministic input(s) $u(k)$. Given the assumption that new external forecasts are better than updated old forecasts, the old states are reset and the new initial state is fully determined by the new forecast, *i.e.* $x(k^*+1) = u(k^*)$. The output $y(k)$ is the observed weather variable and is a scalar. The system matrices are then defined by

$$A(k) = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix} \qquad k \neq k^*, k \in \mathbb{N} \tag{10}$$

$$A(k) = \mathbf{0} \qquad k = k^*, k \in \mathbb{N} \tag{11}$$

$$C = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \tag{12}$$

The system noise or disturbance input is only relevant at time instant $k = k^*$, i.e. when a new external forecast becomes available. The time-varying system matrices B and G are defined as follows:

$$B(k) = G(k) = I \qquad \forall k = k^*, k \in \mathbb{N} \tag{13}$$

$$B(k) = G(k) = \mathbf{0} \qquad \forall k \neq k^*, k \in \mathbb{N} \tag{14}$$

where $A, B, G \in \mathbb{R}^{(M+1)\times(M+1)}$ and $C \in \mathbb{R}^{1\times(M+1)}$.

The Kalman filter matrices Q and R must still be specified. The measurement noise covariance matrix R may be found from the sensor characteristics. The key problem is how to choose Q, the covariance matrix related to the disturbance inputs or system noise. In previous work this system noise covariance matrix of the short-term forecasts has been determined from historical data by comparing forecasts with observations. The initial covariance matrix $P(0)$ is typically chosen as: $P(0) = 10^6 I$, with I the identity matrix.

2.4 Weather data

Three different weather variables, that are most relevant for greenhouse systems, are studied: temperature, wind speed and global radiation. Data is obtained from January 1, 2002 until June 31, 2002 and January 1, 2003 until June 31, 2003. The origin of forecasts and local measurements are specified by:

short term forecasts The commercial weather agency Weathernews Benelux delivered forecasts for location 'Deelen, The Netherlands' (see figure 1). These data become available every six hours and consist of forecasts from 0 to 31 hours ahead with an hourly interval. These external data are extracted from the GFS model (National Weather Service). The data are possibly adjusted by a meteorologist. The data are made available at 0100, 0700, 1300 and 1900 UTC.

local measurements These data are obtained from a greenhouse in 'Wageningen, The Netherlands' located at about 20 km from 'Deelen' (see figure 1). Measurements were stored with a 2 minute interval. The hourly averages were calculated and used for analysis.

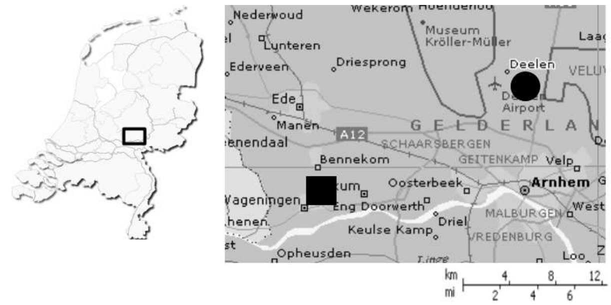

Figure 1. Indication of forecast (●) and measurement (■) location within the Netherlands

3 Results

Forecast data are compared with local observations. For the DBC, matrix Q describes the correlation of the forecast error over a 24 hour horizon. In this experiment it is chosen similar as described by Homleid [5], *i.e.* exponentially decaying until $t + 12$ and then rising again until $t + 23$ ($e^{-0.0744t}$ with $0 \leq t \leq 12, t \in \mathbb{N}$), for each weather variable. The optimal W/V-ratio is found with a line search procedure. Optimality in this case is defined by: minimum average standard deviation (σ) of the forecast error

$$\min_{W/V} \frac{1}{M+1} \Sigma_{i=0}^{M} \sigma_i(W/V) \tag{15}$$

The optimal ratio is calculated over the period January 1, 2002 until June 31, 2002. If W is chosen as 1, and W/V is given, R can be calculated.

The covariance matrix of the forecast error Q of the LAF procedure is calculated for each weather variable with data from January 1, 2002 until June 31, 2002.

As both DBC and LAF are complementary the procedures are run consecutively after a new measurement becomes available. The procedure is run over the period January 1, 2003 until June 31, 2003. The updated forecasts (DBC+LAF) are compared with the original forecasts provided by the weather agency.

The results related to a specific weather input variable contain the optimal W/V-ratio for DBC. In addition, the assumed measurement noise covariance matrix R used in LAF is given. Furthermore, the forecast error, *i.e.* forecast - observation, is calculated and the average forecast error and the standard deviation of the forecast error are presented.

3.1 Temperature

The results are given in figure 2 with a W/V-ratio of 0.011 for DBC and the variance of the measurement noise R in LAF of 0.1 $^\circ C^2$. In figure 2 it can be seen that the bias is reduced for each forecast horizon. The standard deviation is reduced for each forecast horizon but especially up to 10 hours ahead this reduction is clear.

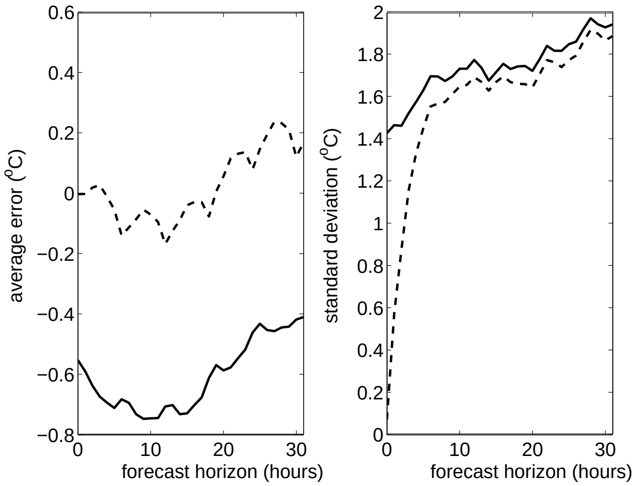

Figure 2. Average forecast error and the standard deviation of forecast error of the original forecasts (—-) and of the adjusted forecasts with DBC+LAF (- - -) of the temperature

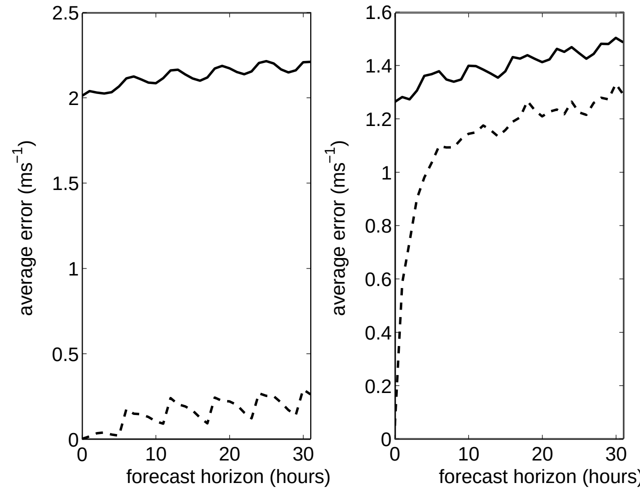

Figure 3. Average forecast error and the standard deviation of forecast error of the original forecasts (—-) and of the adjusted forecasts with DBC+LAF (- - -) of the wind speed

3.2 Wind speed

The W/V-ratio used in the DBC was 0.042. The variance of the observation noise of the wind speed in the local adaptive short term system is assumed to be 0.1 (ms^{-1})^2. The results are summarized in figure 3. The bias is almost completely removed compared to the original forecasts for each forecast horizon. Again, the standard deviation is lowered. The reduction of standard deviation in this case clearly remained until the maximum forecast horizon.

3.3 Global radiation

For global radiation the same LAF procedure as for temperature and wind has been implemented but with $R = 10$ (Js^{-1}m^{-2})^2. However, in DBC the covariance matrix should be adjusted. At night no radiation is available and so no correlation is present. The length of the nights should also vary during the year. For reasons of simplicity the yearly variation is neglected and hence it is assumed that for the whole year no correlation is needed between 1700 and 0500 UTC. The remaining correlations are kept the same as for the temperature case. The optimum W/V-ratio appeared to be around 0.010. In figure 4 the average error and standard deviation for the original forecast and the adjusted forecast are presented. Overall, the bias is reduced. However, only in the first few hours this can be seen clearly. Furthermore, the standard deviation is reduced for each forecast horizon, particularly in the first 5 hours. In addition to this, the peaks are largely removed.

4 Discussion

Bias of the forecast error is often present even if the forecast is related to a specific meteorological measurement location. Therefore, it can be expected that for local measurements, as in agricultural systems, bias is present and is probably larger than for the meteorological station from which the forecast originated. In figures 2 - 4 the bias is clearly present.

When considerable biases are present the DBC works quite well for bias reduction as can be seen in figures 2 and 3. The W/V-ratio appears to be crucial for the performance of DBC. If the W/V-ratio is chosen too large then a loss in performance is observed, *i.e.* the standard deviation increases for larger forecast horizons. For instance, values chosen by [5], *i.e.* $W/V = 0.06$, does not satisfy for temperature in our case. Too low values will lead to neglegible changes in the forecasts. Furthermore, the value for this ratio depends on the weather type and change in weather conditions. A time-varying ratio was proposed by [6]. The problem remains, however, because the value is always obtained from past data.

The line search procedure with optimality criterion (15) to minimize the standard deviation of the forecast error will result in an average standard deviation equal to or lower than the average standard deviation of the original forecast. The bias, however, is then not necessarily minimized. Therefor, other minimization criteria, such as minimum mean square error or minimum bias, must be used according to the defined purpose. It should be noted that the calculated optimal W/V-ratio is kept constant for each year. In practice, the optimal W/V-ratio changes every year. From the results in figures 2-4, however, it can be seen that the calculated optimal ratio is applicable for the following year.

The correlation in the covariance matrix Q in DBC has to be determined for every weather variable. For global radiation, this matrix should also depend on the time of the year because daylight duration can change largely within a

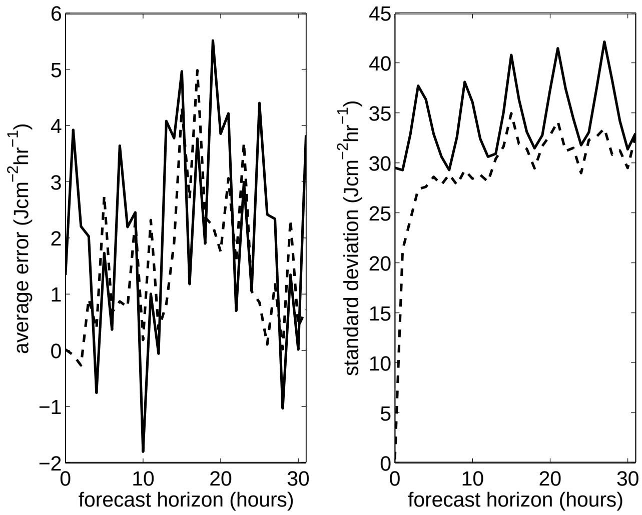

Figure 4. Average forecast error and the standard deviation of forecast error of the original forecasts (—-) and of the adjusted forecasts with DBC+LAF (- - -) of the global radiation

year. In this case the matrix Q must be time-varying.

The LAF procedure largely depends on the choice of covariance matrices Q and R. The matrix Q can be found for every weather variable at a specific location by time series analysis. The period over which the covariance matrices are determined may play an important role. In our study we defined the covariance matrices over a period of six months. This was the same period of the year over which the Kalman filter was run. When seasonal effects on variability of the forecast error are suspected one can choose to determine Q and R from the seasons of a previous year, *e.g.* month by month, quarter by quarter etc. Apart from the seasonal effects, meteorological models may have a more significant effect on the variability of the forecast error. Consequently, one could consider to use a ”window” for the covariances.

In this paper a low variance is assumed for the local measurement. This might be true for the measurement device itself but the variance also depends on how and where the device is installed. The measurement should represent the weather disturbance input of the real system under study. For local measurements devices it is crucial that they are properly calibrated and maintained. Furthermore, the measurement device must be situated on a proper place, *e.g.* a temperature device should not be exposed to direct sunlight. As ambient measurements are frequently used to control greenhouse climates, it is acceptable to update forecasts with local measurements. On the other hand, if it is known that local measurements are unreliable, it is worthwhile to investigate if the measurements can be updated with weather forecasts to generate more reliable measurements. As an example the observations of the meteorological stations of ’Deelen’ and ‘Wageningen’ and the local observations of ‘Wageningen’ are compared in table 1. The meteorological station ‘Wageningen’ is located at about 4 km of the local observations. It can be seen in this table that all differences are quite similar. Intuitively, the difference between local measurements and ‘Wageningen’ meteorological station is expected to be smaller than the others because of the small distance. Apparently, local measurements, as stated in the introduction, can behave quite differently than meteorological measurements.

Table 1. Mean and standard deviation (σ) of observation differences of meteorological stations in ‘Deelen’(D) and ‘Wageningen’(W) and local observations in ‘Wageningen’(L).

	temperature		wind speed		global radiation	
	mean	σ	mean	σ	mean	σ
L-W	0.23	0.91	-0.88	0.84	-1.90	27.5
L-D	0.44	0.92	-1.77	1.17	2.1	29.5
W-D	0.22	1.17	-0.90	1.04	4.5	23.1

In this study DBC and LAF were executed consecutively when a new measurement became available. It could be worthwhile to examine the possibility of integrating both procedures into a single system. For instance, the DBC supposes a full correlation of the diurnal pattern, *i.e.* a certain error obtained now will also be present tomorrow. A more valid assumption is that this relation is exponentially decaying. As a result the W/V ratio can be chosen larger. As an alternative to the stochastic filtering approach one may also consider an unknown-but-bounded error approach [see *e.g* 9, 10].

5 Conclusions

It has been demonstrated in this paper that both bias and standard deviations of forecast errors are reduced for three different weather variables: temperature, wind speed and global radiation. The tuning parameters of DBC and LAF, however, must be chosen carefully. Using historical for tuning the parameters gives adequate results.

Acknowledgements

We are indebted to Agrotechnology an Food Innovations and WNI for providing the data. This project (weather in control II) is supported with a grant of the Dutch Programme EET (Economy, Ecology, Technology) a joint initiative of the Ministries of Economic Affairs, Education, Culture and Sciences, and of Housing, Spatial Planning and Environment, in co-operation with PRIVA BV, Tolsma, WNI and Agrotechnology and Food Innovations. The programme is run by the EET Programme Office, a partnership of Senter and Novem. For details of the project we refer to the website: http://www.atoapps.nl/Wic2

References

[1] Z. S. Chalabi, B. J. Bailey & D. J. Wilkinson, A real-time optimal control algorithm for greenhouse heating, *Computers and Electronics in Agriculture, 15* (1), 1996, 1–13.

[2] K. J. Keesman, D. Peters & L. J. S. Lukasse. Optimal climate control of a storage facility using local weather forecasts. *Control Engineering Practice, 11* (5), 2003, 505–516.

[3] R. F. Tap, L. G. Van Willigenburg & G. Van Straten. Receding horizon optimal control of greenhouse climate using the lazy man weather prediction, *Proc. of the 13th IFAC World Congress*, San Francisco, USA, 1996, 387–392.

[4] T. G. Doeswijk and K. J. Keesman, Adaptive weather forecasting using local meteorological information, *Biosystems Engineering*, 2005. Submitted.

[5] M. Homleid. Diurnal corrections of short-term surface temperature forecasts using the kalman filter, *Weather and Forecasting, 10*(4), 1995, 689–707.

[6] G. Galanis and M. Anadranistakis. A one-dimensional kalman filter for the correction of near surface temperature forecasts, *Meteorological Applications, 9*, 2002, 437–441.

[7] World Meteorological Organization (WMO). *Guide to Meteorological Instruments and Methods of Observation, volume 8* (Geneva, sixth edition: WMO, 1996).

[8] A. Gelb. *Applied Optimal Estimation* (Cambridge, Massachusetts: MIT Press, 1974).

[9] F. C. Schweppe. *Uncertain Dynamic Systems* (Englewood cliffs, New Jersey: Prentice Hall, 1973).

[10] K. J. Keesman. Weighted least-squares set estimation from l_∞-norm bounded-noise data, *IEEE Trans. Autom. Control*, 42(10), 1997, 1456–1459.