DANISH METEOROLOGICAL INSTITUTE

TECHNICAL REPORT

00-04

Kalman Filtering of DACFOS
First Verification Report
Period July - December 1998

Jérôme Chenevez

DMI

COPENHAGEN 2000

ISSN: 0906-897X (printed version)
1399-1388 (online version)

LIST OF CONTENTS

1. Introduction

The Kalman filtering of DACFOS for ozone forecasts at the DMI monitoring station in Jægersborg became fully operational on the 1st of July 1998. Since this date 48-hour forecasts have been presented four times daily on the DMI public Internet homepage, together with others DACFOS ozone forecasts. The forecasts are presented in two forms: 1) at the location of Jægersborg, corresponding to the region StorKøbenhavn on the map of Denmark is displayed a colour indicating the maximum ozone concentration level expected for the present day and, on another picture, for the next day ; 2) on a third picture is displayed a graph showing the evolution of the forecast ozone concentration for the next 48 hours. This graph is actually obtained by averaging four distinct Kalman filters of DACFOS ; details about the system are given in DMI Technical Report 98-15.
In the present report results verification for the operational Kalman filter for ozone forecasts in Jægersborg is presented for the 2nd semester of 1998.

2. Results

In order to test the quality of the Kalman filter forecasts the following verification compares model results with observations (i.e., ozone concentration measurements in Jægersborg), but also with DACFOS ozone forecasts, allowing to verify the suitability of using the Kalman filter. In the first subsection, observed and forecast maximum daily ozone concentrations during the last 6 months of 1998 are compared. In the next subsection a verification of the performances of the "colour-forecast" (forecast of the class of the daily maximum ozone concentration) is presented. In the last subsection some statistics are done on all the 4-daily forecasts performed during the 2nd semester of 1998. Finally some conclusions are given in section 3.

2.1 Monthly plots of maximum ozone concentration

Plots of the maximum ozone concentration observed in Jægersborg and forecast (both by DACFOS and the Kalman filter) each day at 00H (UTC) are shown in Annexe 1 for each month of the 2nd semester 1998. During all the period, only 12 days are missing because lack of data ; the hole on July 9th is due to missing observations. One can see on these plots that during the whole period, the threshold of 60 ppb (120 $\mu g/m^3$, corresponding to the high level of ozone concentration) is never reached and the measured ozone concentrations are always low or medium.
Figure 1 displays monthly means of 1-hour averaged ozone concentration observations in Jægersborg between July and December 1998. This shows that the ozone season (Summer period) in 1998 was not much higher than during Fall and rather low compared to the year 1997, for which the average value during July and August was 29 ppb.

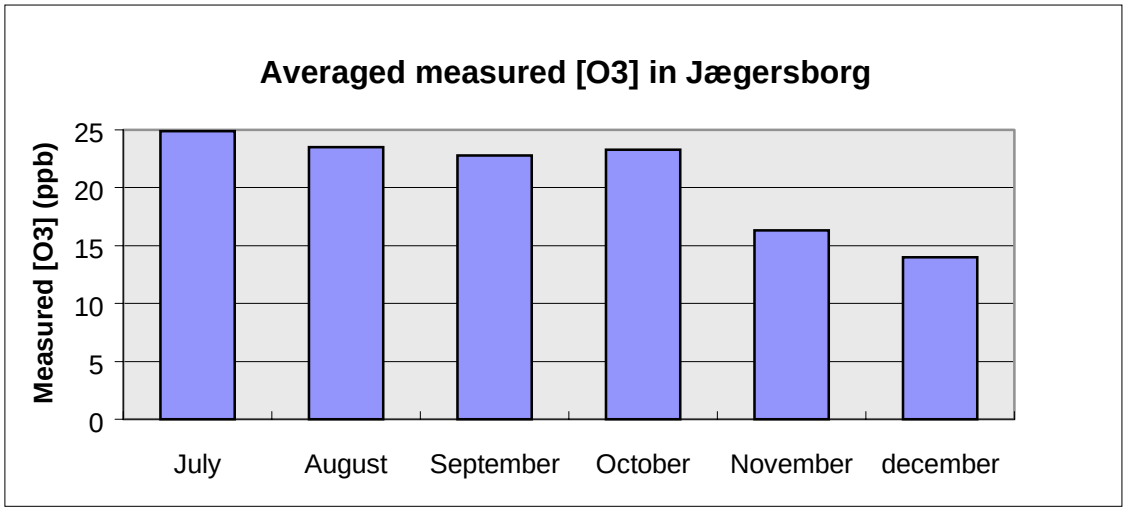


Figure 1 : Monthly means of 1-hour averaged ozone concentrations in Jægersborg.

As shown on the first figure of Annexe 1, the highest measured value in 1998 was 59,5 ppb on July 21st, which was the warmest day of the year with 27 °C measured in Jægersborg. This peak was well forecast by the Kalman filter (57,2 ppb), while DACFOS overestimated it at 78 ppb. Figure 2 presents the 48-hour forecast performed on July 21st at midnight and shows that the anticipated time of the maximum value also was pretty good both for DACFOS and the Kalman filter, as well as the relative lower value of the second day.

For the rest of the period, only a few days in August were over 50 ppb but for which DACFOS performed better than the Kalman filter, which underestimated the peaks. Generally, over the whole period, the forecast maximum values were too low, but the measured concentrations being however not high the differences were not so big. There were nevertheless many days, especially in October, where the forecast (both by DACFOS and the Kalman filter) class was Low (under 30 ppb) though the observed class was Medium (over 30 ppb).

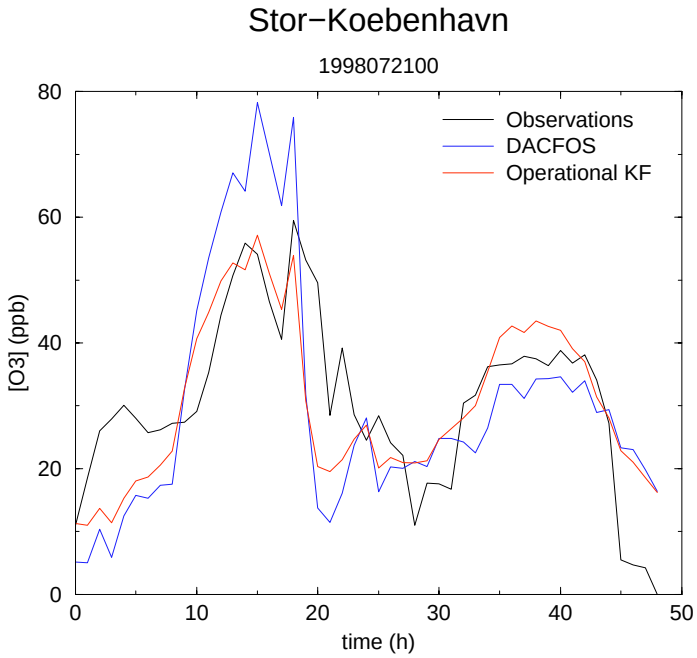


Figure 2 : Ozone concentration forecasts and observations in Jægersborg between July 21st at 00h and 23rd at 00h.

2.2 Performances of the class-forecasts

Forecast of the class of the ozone concentration is given daily by a colour indicating the maximum ozone concentration value reached each day. These classes are called LOW (green) when the maximum ozone concentration is under 30 ppb, MEDIUM (orange) when it is between 30 ppb and 60 ppb, HIGH (yellow) between 60 ppb and 90 ppb, and VERY HIGH (red) when it is over 90 ppb. Figure 3 shows the number of days observed and forecast in each class during the 2nd semester of 1998. Of course, as previously mentioned, only the two first classes are occupied, but by DACFOS, which has a forecast in the HIGH class (July 21st).

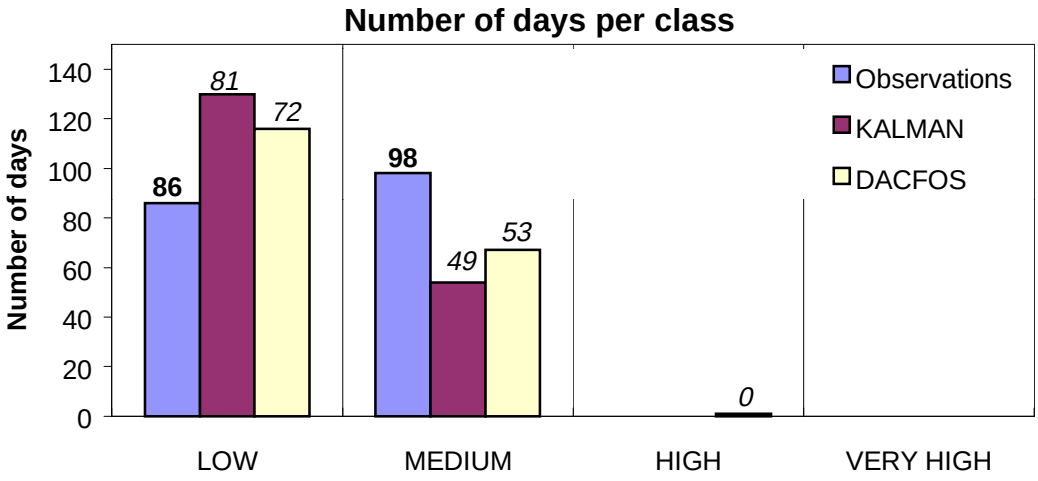


Figure 3 : Number of days during the 2nd semester 1998 in each of the four ozone concentration classes for measurements in Jægersborg and both DACFOS and Kalman filter forecasts at 00H (UTC). The *italic* numbers are the numbers of days where the forecast class coincides to the observed class.

On the 86 days observed in the LOW class, the Kalman filter predicted 81 of them for a total number of 130 forecast days in this class, which yields an efficacy of the forecast of 62,3%, while DACFOS predicted 72 correct days but only with 116 forecast days in this class, yielding also an efficacy of 62,1%. There were 98 days in the MEDIUM class, on which the Kalman filter forecast correctly only 49 days among 54 forecast days in this class (efficacy: 90,7%), while DACFOS gave 53 correct days but for 67 forecast days in this class (efficacy : 79,1%).
Let us now have a look at the class performances of the whole of the 4-daily forecasts both for the first and the second forecast days ; this is shown on figure 4.

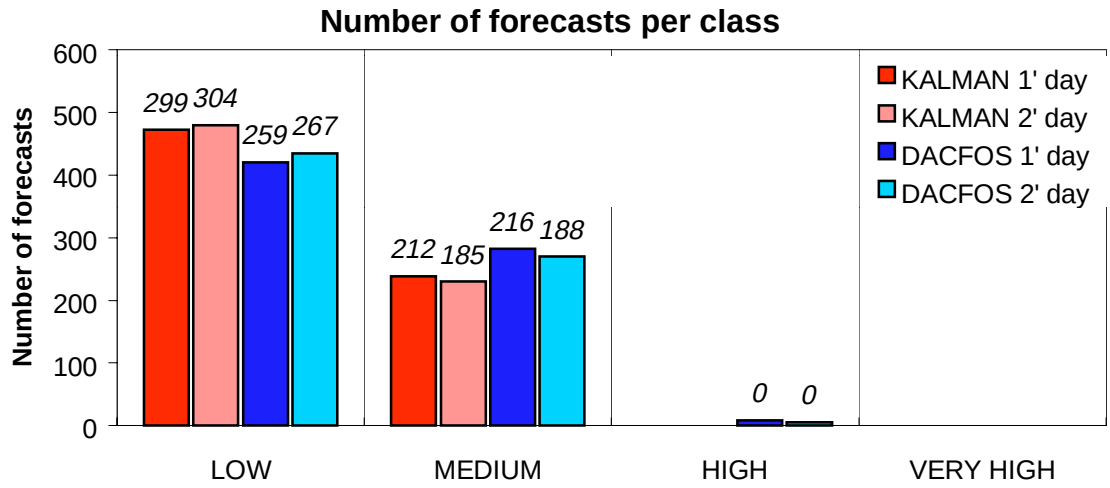


Figure 4 : Total number of 2-days forecasts during the 2nd semester 1998 in each of the four ozone concentration classes for both DACFOS and Kalman filter forecasts. The *italic* numbers are the numbers of days where the forecast class coincides to the observed class. There are in all 710 forecasts.

From the results of figure 4, one can calculate the total efficacy of the forecasts; for the first day 72% of the Kalman filter class-forecasts coincided with the observed class (63% for the LOW class and 89% for the MEDIUM class) against 67% of DACFOS class-forecasts (62% for LOW, 77% for MEDIUM); for the second day these rates were 69% for the Kalman filter (63% for LOW, 80% for MEDIUM) and 64% for DACFOS (61% for LOW, 70% for MEDIUM).
In figure 5 the distribution of the errors between DACFOS and Kalman filter forecast daily maximum concentration and measured daily maximum concentration are shown for the first and the second forecast day. The figure shows that both models have a slight tendency to underestimate the peaks, but this tendency is lower for the Kalman filter, especially on the first day, than for DACFOS.

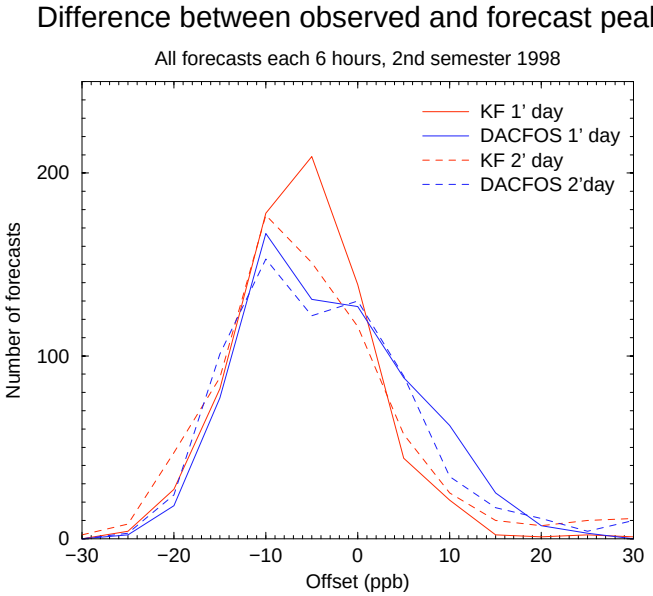


Figure 5 : Histogram of the difference between forecast and measured daily maximum ozone concentration in the 2nd semester of 1998 in Jægersborg.

2.3 Statistics

Some further statistics are now presented that test the accuracy of each hour of the forecasts performed four times daily during the 2nd semester of 1998; there are 710 forecasts available representing more than 15.000 points.

Figure 6 shows a scatter plot between 1 hour averaged ozone concentration measured in Jægersborg and model ozone concentration both for the Kalman filter and DACFOS. It is obvious that the Kalman filter has a larger correlation (r=0,64) than DACFOS (r=0,47).

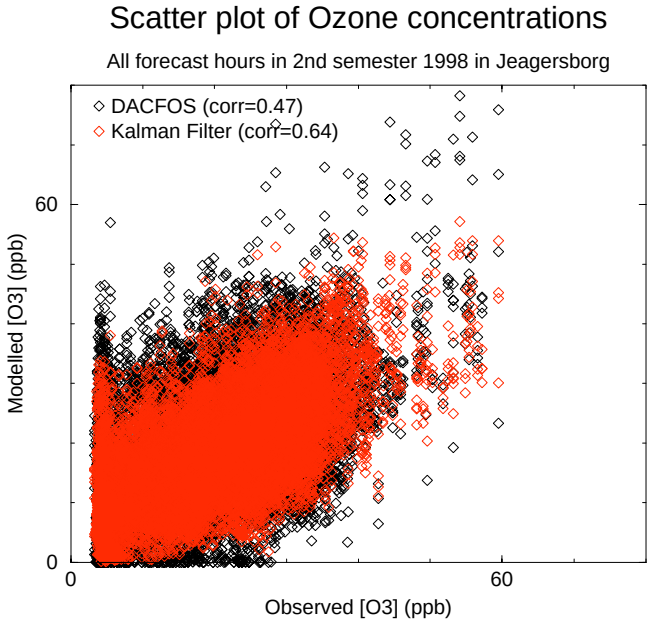


Figure 6 : Scatter plot for DACFOS and Kalman filter hourly ozone concentrations between July 1st and December 31st 1998 in Jægersborg.

In figure 7 the evolution of the correlation is shown as a function of the forecast time both for DACFOS and Kalman filter, as well as for the persistence forecast. This shows again that the Kalman filter has the highest correlation coefficient all over the forecast period.

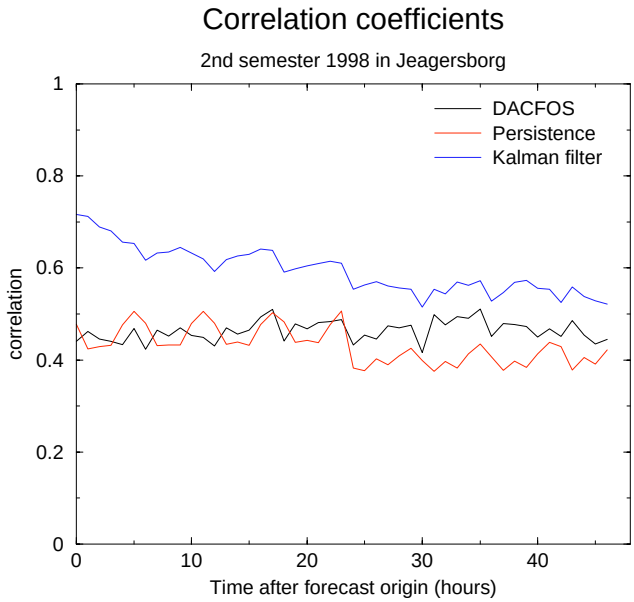


Figure 7 : Comparative evolution of the correlation coefficients for DACFOS, Kalman filter and the persistence forecasts during time after forecast origin.

Figure 8 presents the evolution of the forecast error with time after forecast origin. This time the persistence has the lowest error and the Kalman filter is contained between the persistence and DACFOS has smaller bias than the Kalman filter during the 36 first hours.

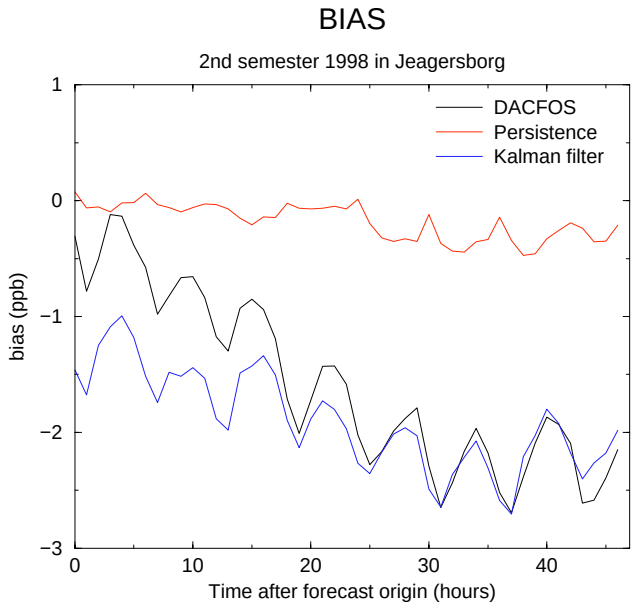


Figure 8 : Comparative evolution of the biases during time after forecast origin for DACFOS, Kalman filter and the persistence forecasts.

In the same way figure 9 shows that the Kalman filter has always the lowest rms value, but this slightly increases with time after forecast origin.

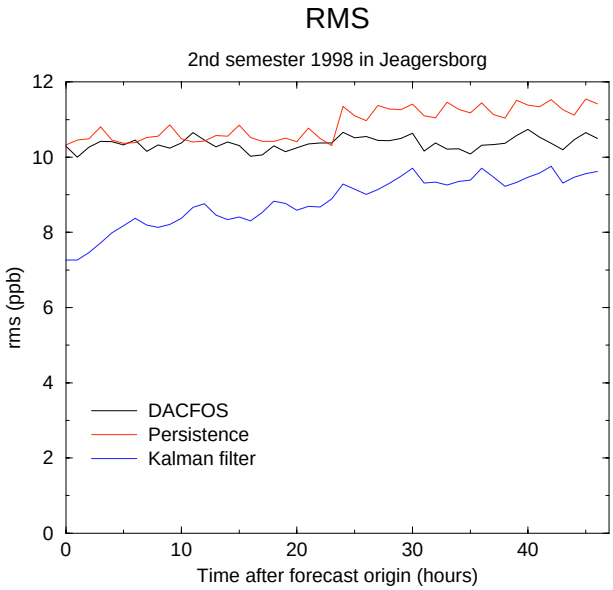


Figure 9 : Comparative evolution during time after forecast origin of the rms values for DACFOS, Kalman filter and the persistence forecasts.

Table 1, by combining the results of the four daily forecasts between July and December 1998, summarises the statistics obtained for the four distinct Kalman filters that use different meteorological variables and which are averaged to obtained the operational Kalman filter used for the Internet forecasts. One can read in this table the statistic performances of each of these four Kalman filters, as well as those of the operational (averaged) Kalman filter, compared with those of DACFOS and the Persistence forecasts.
In this table the "Total Averages" are the mean values for the forecasts over all the 2nd semester o 1998.
<err> is the mean error or bias.
<|err|> is the mean absolute error.
rms is the root-mean-square of the error.
corr is the correlation.
There are also statistics of the maximum values (Peak) of the 1st and the 2nd forecast days: <err> is the bias of the forecast of the day's peak and rms is the root-mean-square of the error on the forecast of the day's peak.
One can see that all the four Kalman filters have higher correlation and lower rms and mean absolute error values than both DACFOS and the Persistence ; the averaged Kalman filter has the highest correlation and the lowest rms and mean absolute error values. This shows that using the average of the four Kalman filters slightly ameliorates the forecasts. Nevertheless, both DACFOS and the Persistence have lower biases and rms values for the peaks forecasts. This means that during the 2nd semester of 1998 the Kalman filter was the best to follow the behaviour of the ozone concentration, but not the best to predict the daily maximum value.

Summary of Performances of Dacfos and Kalman Filters for Ozone-Concentration Forecasts.
Reading 710 logfile(s) each with 4 filters
Kalman Filter Parameter Values:
Q-value = 0.0001 R-value = 1000 Memory = 200
u1: Use Dacfos mean
u2: Use Dacfos 5 levels
u3: Use Hirlam Temperature
u4: Use Hirlam Wind
u5: Use HIRLAM ABL (Q-value = 0.00001 R-value = 10000 Memory = 400)
u6: Use HIRLAM Surface Heat flux

****

JÆGERSBORG 2nd semester 1998 (4 forecasts per day)

Statistics for the Observations:

Mean	St.Dev.	<Peak>	Peak-stdev.
20.68	10.01	30.41	8.94

Mean Errors and Root Mean Square Errors for Forecasts and Peak Forecasts:

	Total Averages				Peak 1st day		Peak 2nd day	
	<err>	<|err|>	rms	corr	<err>	rms	<err>	rms
Dacfos:	-1.56	8.33	10.36	0.46	-1.49	8.87	-2.69	9.08
Persistence:	-0.18	8.47	10.89	0.43	0.00	7.75	0.19	9.11
Kalman Filter using:								
u1 u2 u3 u4 u5 u6								
1 1 1 0 0 0	-2.48	7.23	8.96	0.58	-4.95	8.42	-4.94	9.10
1 1 1 1 0 0	-2.51	7.13	8.88	0.58	-6.03	8.98	-6.56	10.01
1 1 1 1 0 1	-2.64	7.56	9.42	0.56	-3.90	8.45	-4.36	9.76
0 1 1 1 1 0	-0.58	7.36	9.47	0.56	-2.56	8.34	-2.77	10.34
Averaged KF	-1.89	7.05	8.85	0.59	-4.56	8.16	-5.11	9.54

Table 1 : Detailed statistics for the four Kalman filters used in the operational Kalman filter setup, compared with DACFOS and Persistence forecasts. Each distinct Kalman filter is characterised by three intern parameters (Q, R and Memory) and a set of extern variables (u1-6). The operational Kalman filter used for Internet forecasts is the arithmetic average of these four Kalman filters. All values except the correlation (corr) are in ppb.

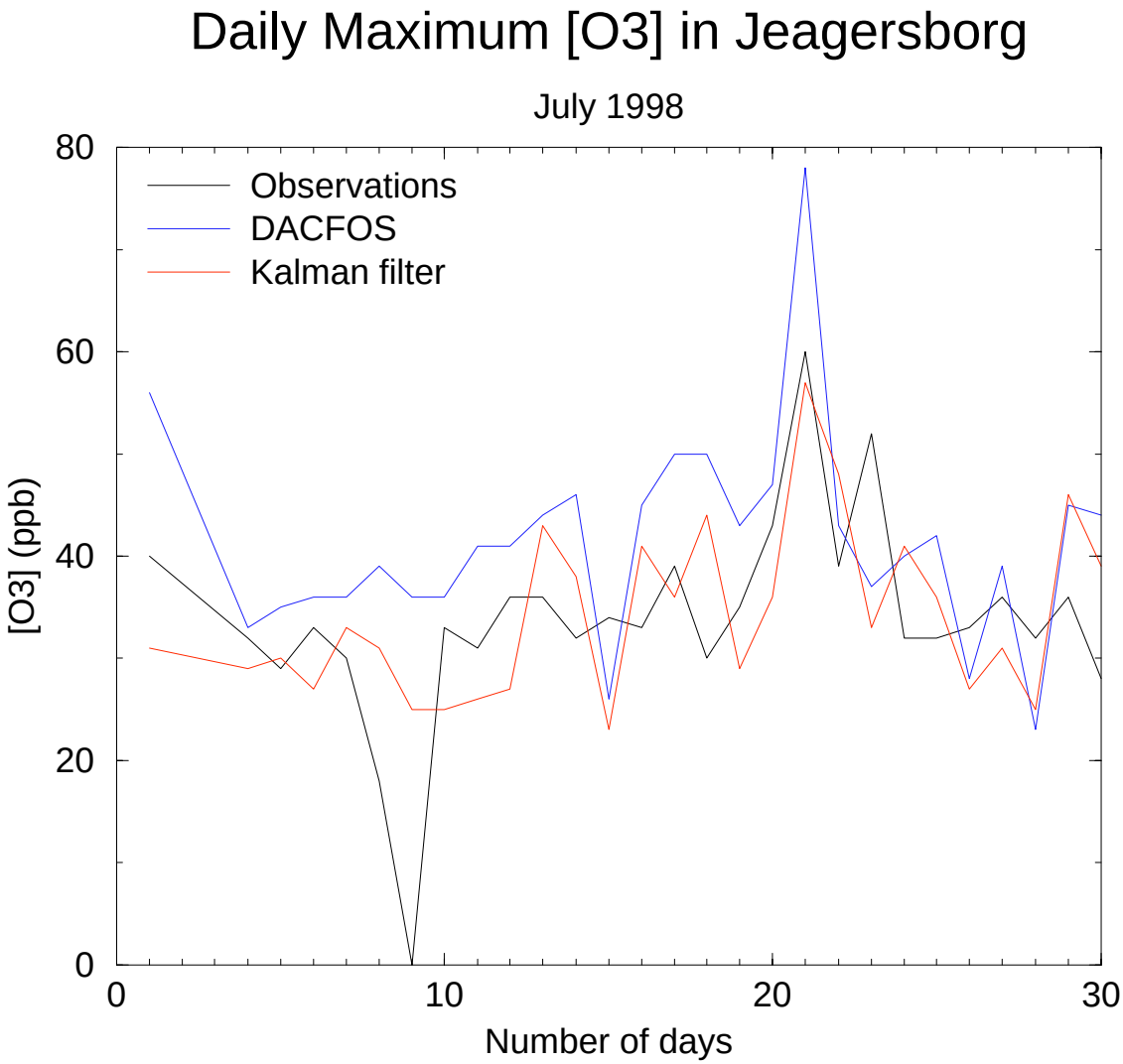
3. Conclusion

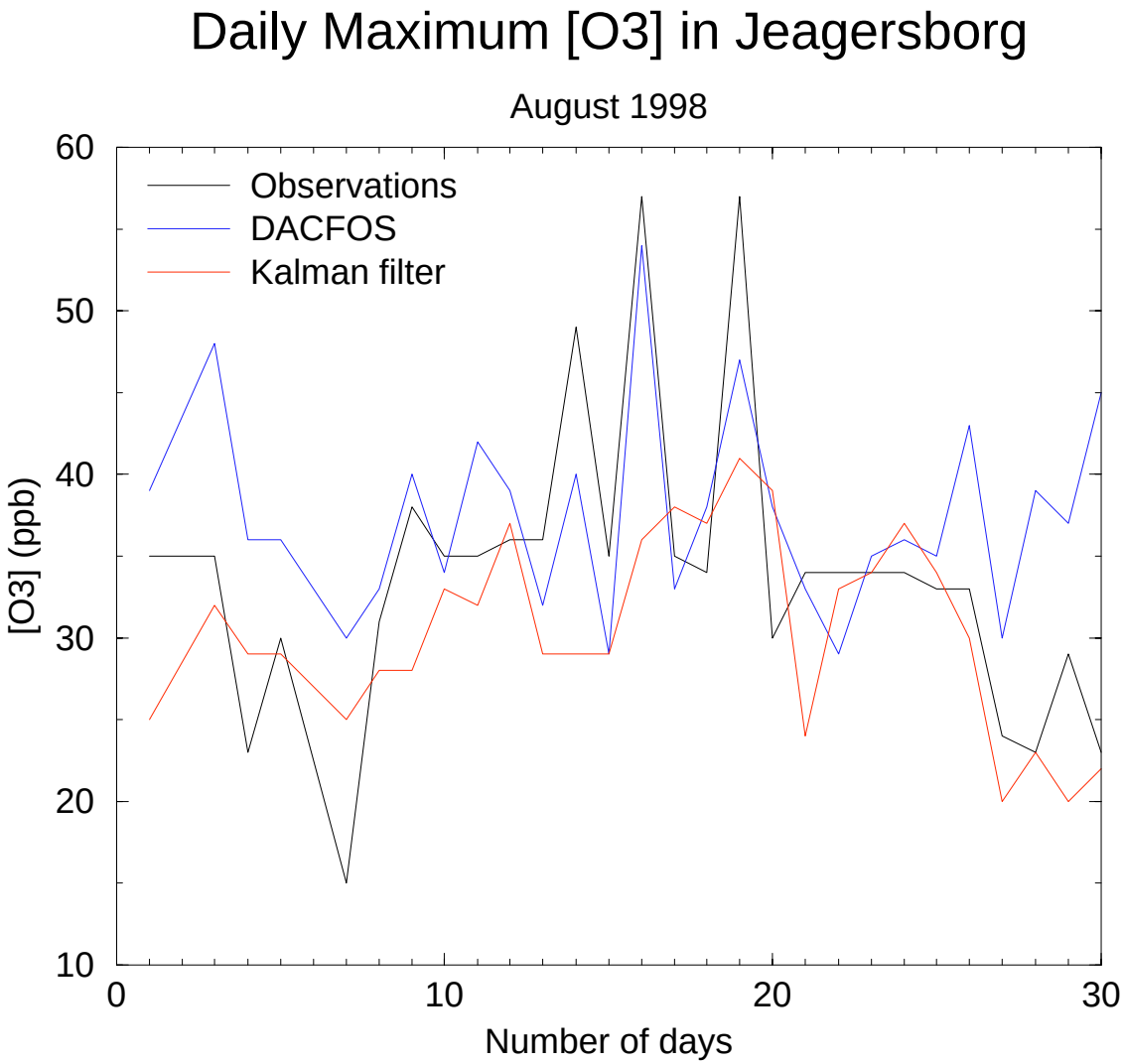
We have seen that the quality of the operational Kalman filter forecasts for ozone concentration in Jægersborg during the 2nd semester of 1998 was generally satisfying. Indeed, by looking globally at the performance of the Kalman filter one can see that most of the daily ozone levels were correctly predicted and the forecasts were the most often better than those of DACFOS. However, there were also some days when the maximum ozone concentration was better predicted by DACFOS than by the Kalman filter, especially when the peak was in the MEDIUM ozone level class ; but this may be explained by the fact that most of DACFOS forecasts were in the MEDIUM class, which has shown to yield a lower efficacy of DACFOS compared to the Kalman filter. Nevertheless, as mentioned in the beginning of this report, the year 1998 in Jægersborg was rather a “low-ozone-concentration-year” compared to the previous years ; this mostly was due to a relatively cold and sunless summer in Denmark. This partly explains why the Kalman filter did not perform better, though it uses to be the best predictor of the worse ozone pollution days that are, the most often, experienced if the

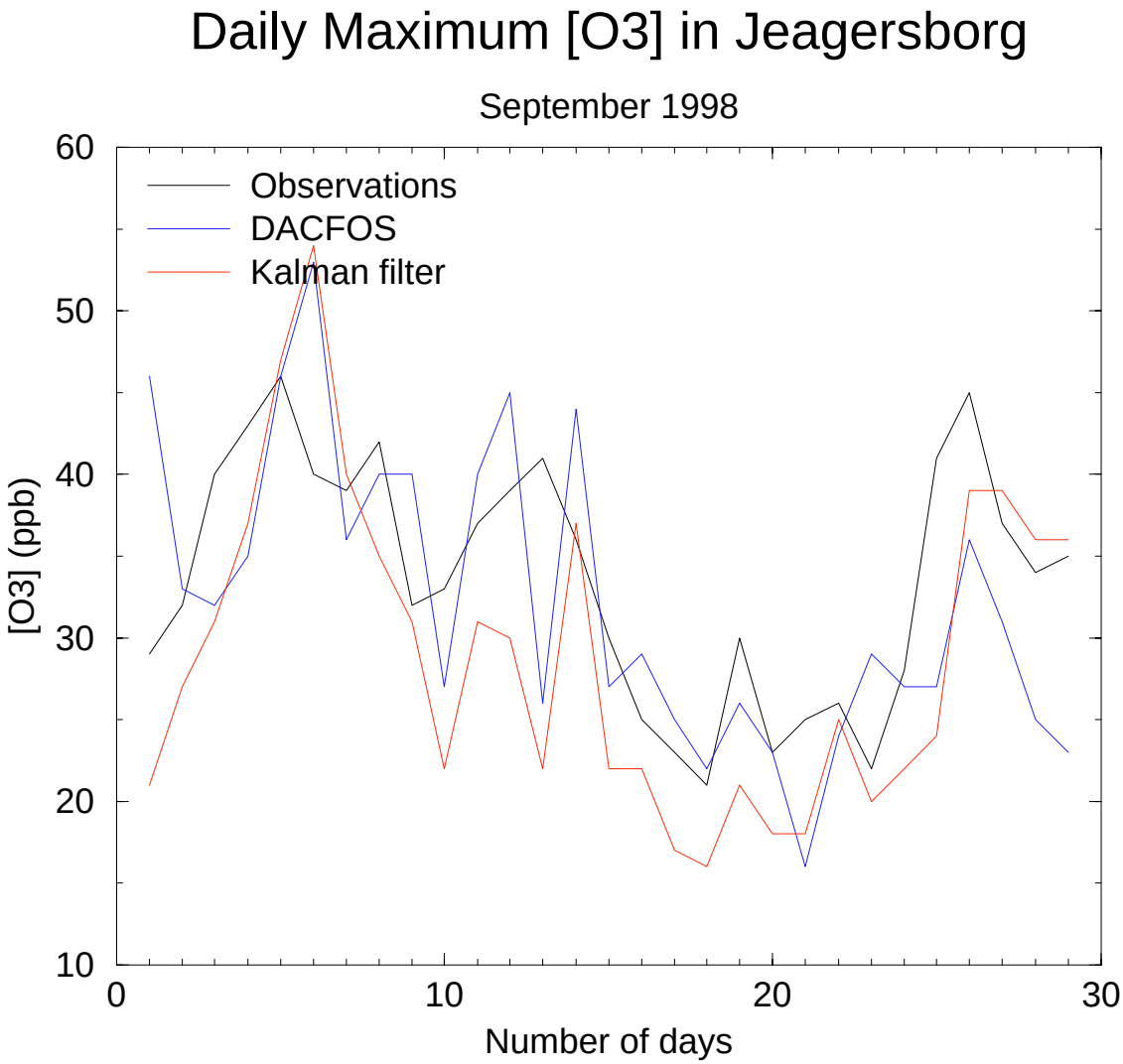
meteorological conditions are suitable to yield rather high ozone concentration values, as for example on 21st of July 1998.
In conclusion, comparing the global performances of DACFOS and of the Kalman filter of DACFOS, it has been showed that 1) using a Kalman filter improves the forecasts and 2) averaging the forecasts of four distinct Kalman filters using different meteorological variables also ameliorates the ozone concentration forecasts ; this justifies to continue to use and extend the present operational Kalman filter set-up.

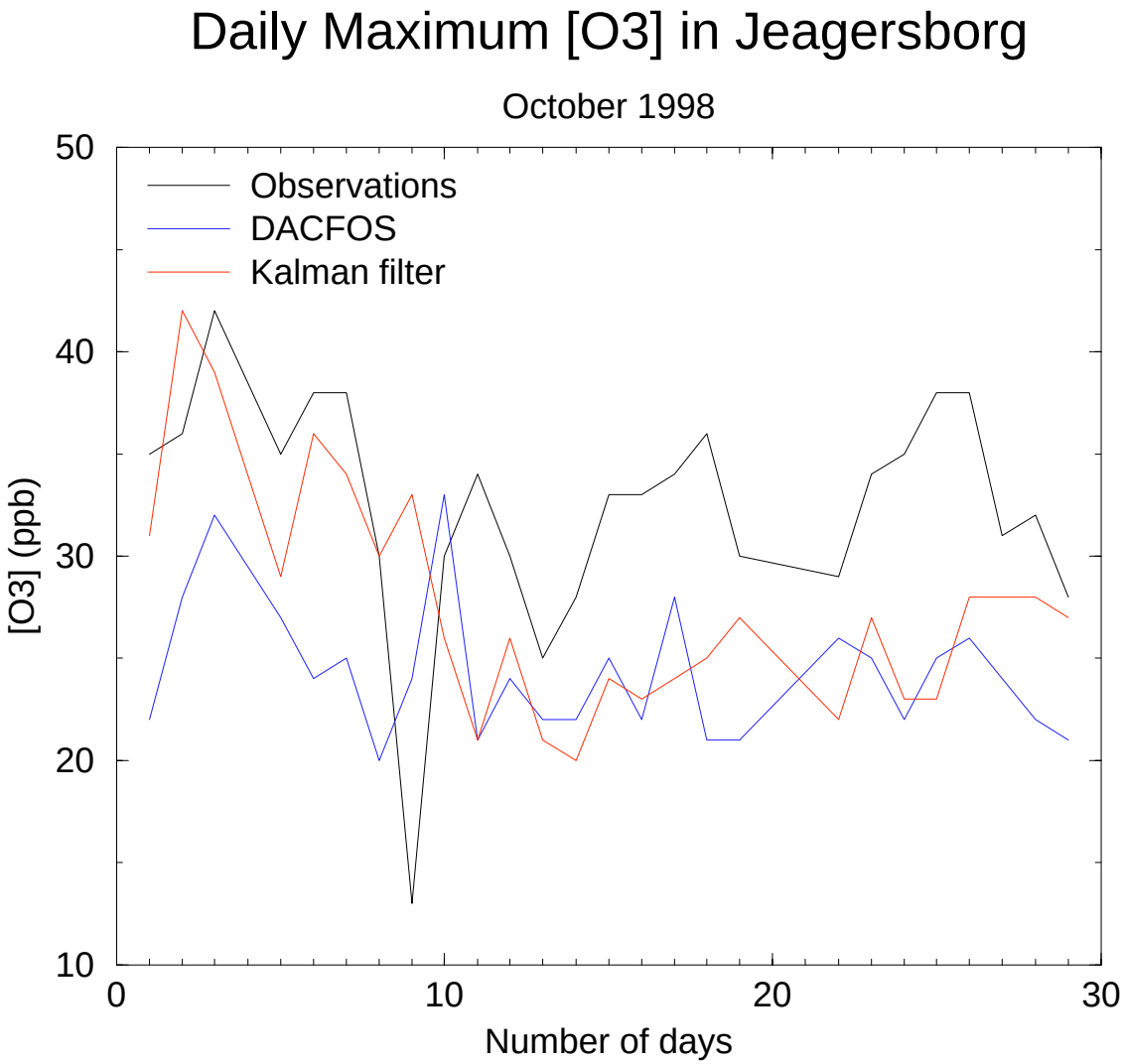
ANNEXE 1

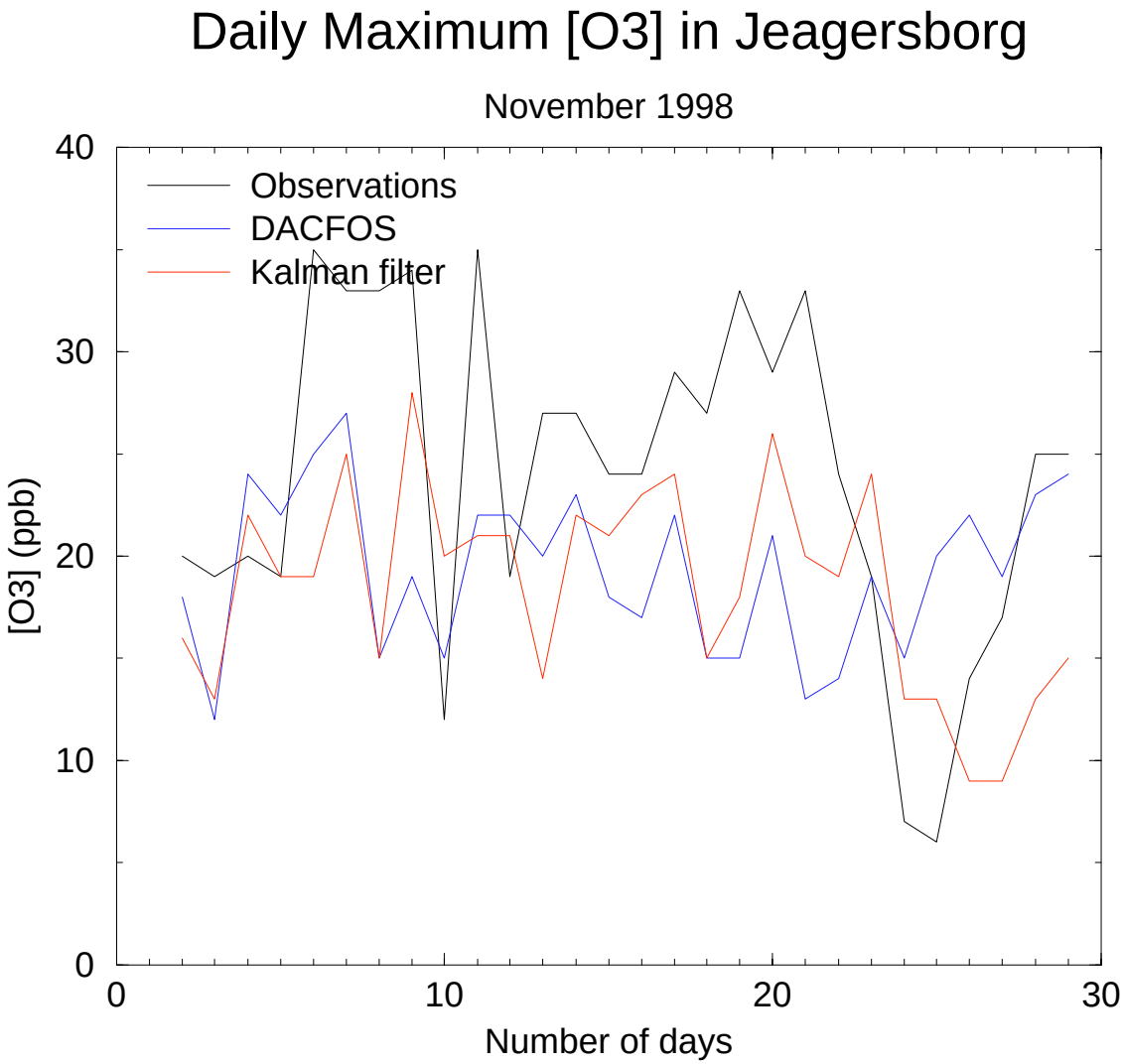
Monthly plots of observed and forecast ozone concentrations in Jægersborg for the time period 01/07/98 to 31/12/98

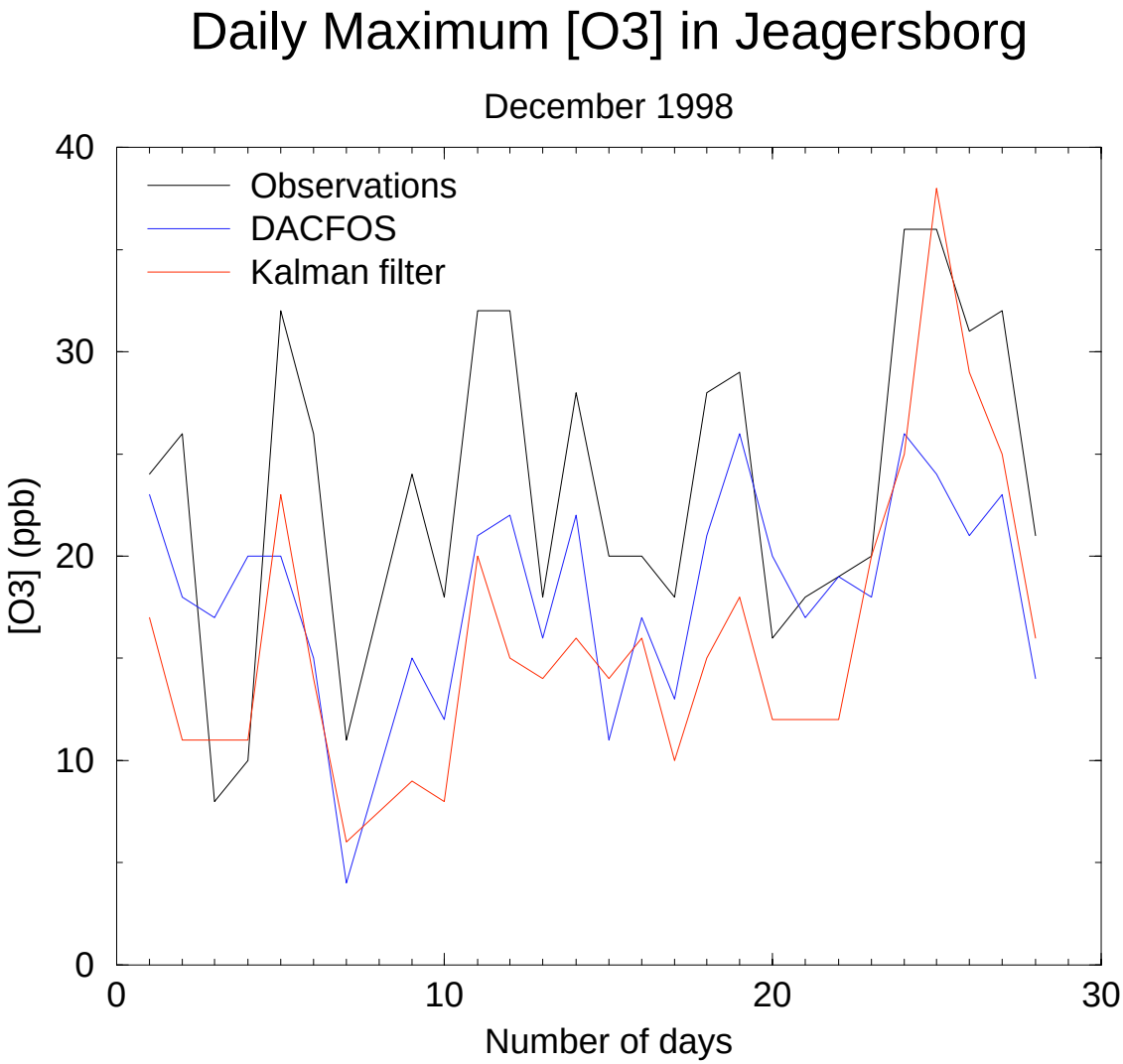

Daily Maximum [O3] in Jeagersborg
July 1998
Observations
DACFOS
Kalman filter
[O3] (ppb)
Number of days
0
20
40
60
80
0
10
20
30


Daily Maximum [O3] in Jeagersborg
August 1998
Observations
DACFOS
Kalman filter
[O3] (ppb)
Number of days
60
50
40
30
20
10
0
10
20
30


Daily Maximum [O3] in Jeagersborg
September 1998
Observations
DACFOS
Kalman filter
[O3] (ppb)
Number of days
60
50
40
30
20
10
0
10
20
30


Daily Maximum [O3] in Jeagersborg
October 1998
Observations
DACFOS
Kalman filter
[O3] (ppb)
Number of days
50
40
30
20
10
0
10
20
30


Daily Maximum [O3] in Jeagersborg
November 1998
Observations
DACFOS
Kalman filter
[O3] (ppb)
Number of days
0
10
20
30
40


Daily Maximum [O3] in Jeagersborg
December 1998
Observations
DACFOS
Kalman filter
[O3] (ppb)
Number of days
0
10
20
30
40