Assimilation of Mode-S EHS observations in ALADIN BlendVar

Benedikt Strajnar, Slovenian Environment Agency

contributions by Alena Trojáková, Patrik Benáček, CHMI

1   Introduction

Mode-S Enhanced Surveillance (EHS) observations (de Haan, 2011) are a relatively new and promising high-resolution aircraft observations. Data are currently collected over large part of western and in some parts of central Europe and then preprocessed and distributed by KNMI.

EHS observations were recently also made available in the OPLACE preprocessing system in the OBSOUL format suited for data assimilation. This study aims at understanding the observation properties in terms of quality compared to standard AMDAR aircraft observations, and the observation impact in the ALADIN forecast. This is to be compared with earlier evaluation within HIRLAM community (de Haan and Stoffelen, 2012). With much increased temporal and spatial resolution compared to AMDAR, the performance of preprocessing, the observation error correlation and the need to thin the data accordingly is to be assessed. Based on the evaluated impact on Czech BlendVar assimilation system we would like to obtain some recommendations for Mode-S EHS usage within LACE.

2   Quality

The quality of Mode-S EHS observations can be estimated by comparison to other observations close in space and time, as done previously. In this approach one is limited to rather sparse radiosonde or AMDAR observations. On the other hand a comparison with NWP offers a reference for every Mode-S observations. The OBS minus NWP departure, however, is a sum of observation plus model (forecast) error and must thus be evaluated with care. Rather than providing an observation error magnitude, this method enables inter comparison between different types of aircraft observations.
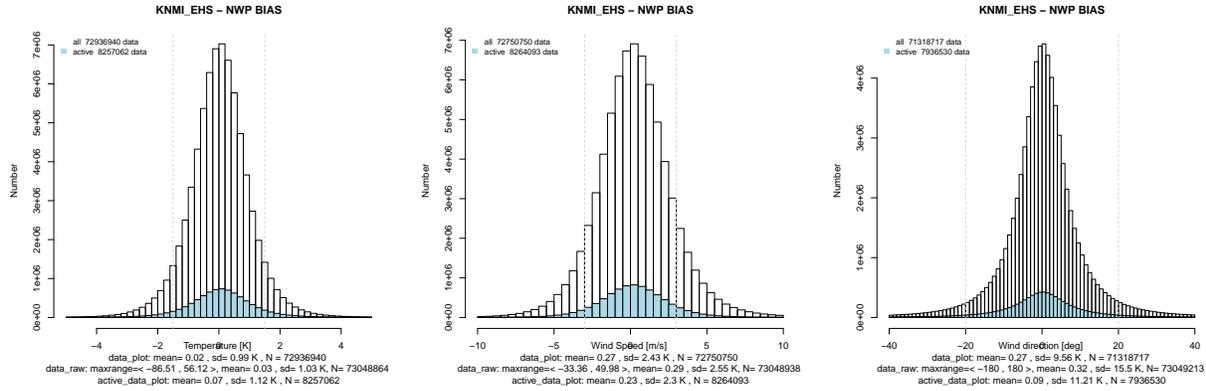
Figure 1: Obs-minus-guess (OMG) departures of 10% randomly sampled EHS data. The vertical lines show thresholds used for whitelisting based on standard deviation. All data (white bars) and data selected by screening (ble bars) are shown.

2.1 Observation minus guess departures

To provide a reliable statistics for comparison, the observation minus guess (OMG) departures with respect to the operational ALADIN model were computed. The period was 10 months (July 2016 - April 2017). OMGs were computed separately for each hour using operational forecast ranges from 7-11 hours. This gave around 730 mio observations for each variable. Due too large data set and to avoid too memory demanding analysis a random sample of all data was analyzed (e.g. 1 or maximally 10% of all EHS observations). The histogram of OMG departures for temperature and two wind components on 73 mio. observations is shown in Fig. 1. The departures are normally distributed. There are some outliers (see figure for maximal range of OMG departures) which need to be rejected. In the further analysis we use certain thresholds to remove outliers which may spoil the statistics.

Vertical profile of temperature OMG departures is presented in Fig. 2. Here the 1% or 10% sampling is compared for one set of whitelisting criteria: aircraft must have absolute bias of less than 0.5 K, less than 1.5 K standard deviation and there must be 1000 and 100 observations in 10% and 1% sampling, respectively. From the number of data or the number of whitelisted aircraft we can conclude that sampling is robust enough to represent the complete EHS data set.

Figures 3 - 5 show comparison of vertical profiles of OMG for all the variables. EHS is compared to AMDAR and Slovenian Mode-S MRAR observations. For EHS all, whitelisted and active (selected by screening) observations are shown. First, we observe very small difference between all and whitelisted observations (the latter have slightly better statistics). The whitelist criteria are given in the next subsection. Second, the active observations have worse statistics which probably comes from much less data in the statistics (rightmost panel). Comparison to AMDAR shows that EHS wind have even better statistics than AMDAR especially near the ground. This might be due to extensive calibration of EHS data. For temperature, EHS has worse standard deviation than AMDAR which is expected given the approximate method to derive EHS temperature (de Haan, 2011). The bias however is very small. The MRAR from Slovenia has somewhat different behavior for temperature bias compared to EHS and AMDAR, and is worse for wind. Here we have to keep in mind that
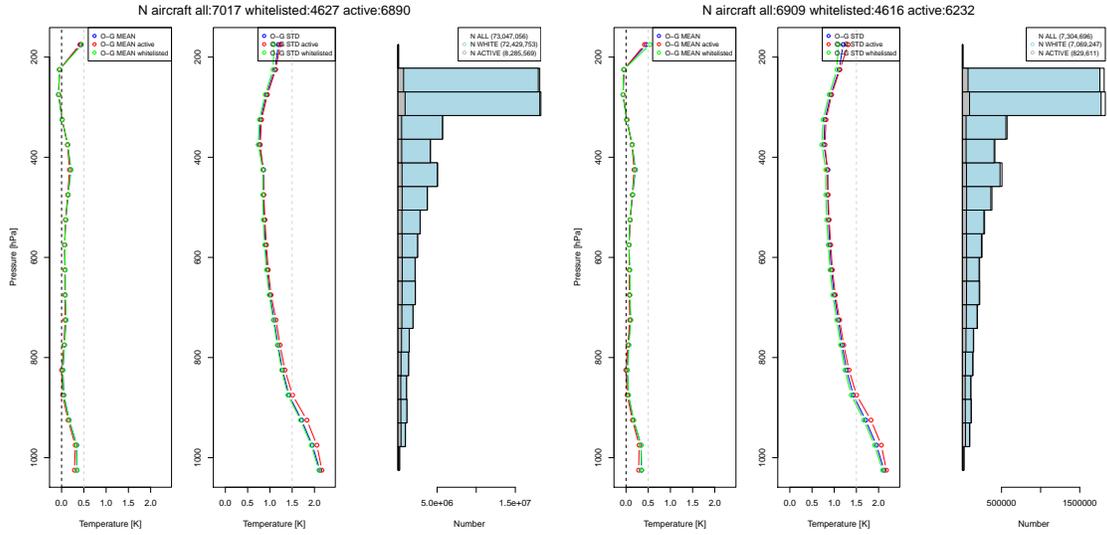
Figure 2: Vertical profile of mean and standard deviation (STD) temperature OMG departures using 10% (left) and 1% (right) randomly sampled EHS data. Shown are all, active (in screening) and whitelisted observations. The requested obs. number for whitelisting are 1000 (right) and 100 (left) and this corresponds to 10000 data in the complete data set.
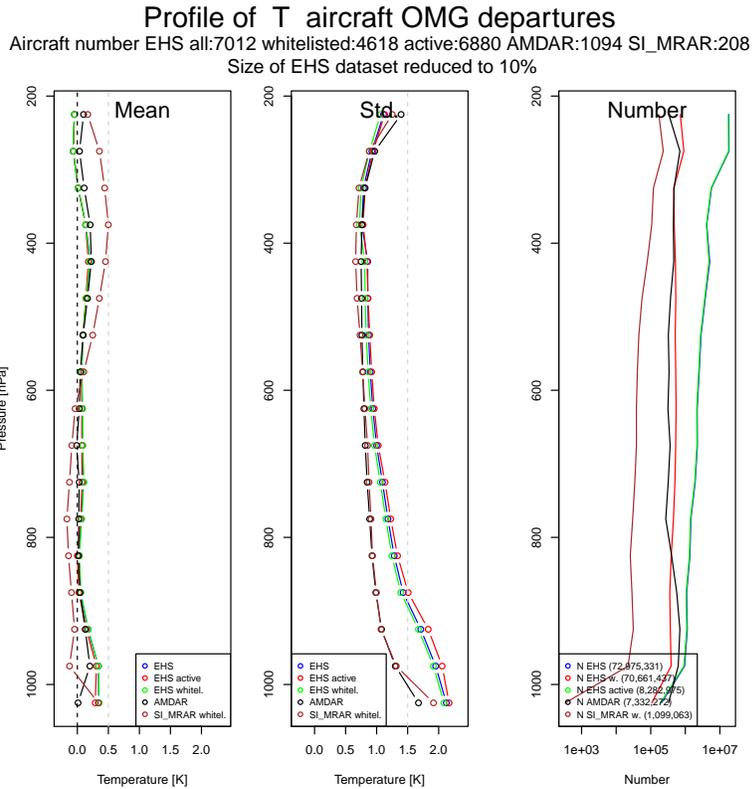

Figure 3: Veritical profile of mean and STD of OMG departures for temperature using 10% of randomly sampled EHS data.
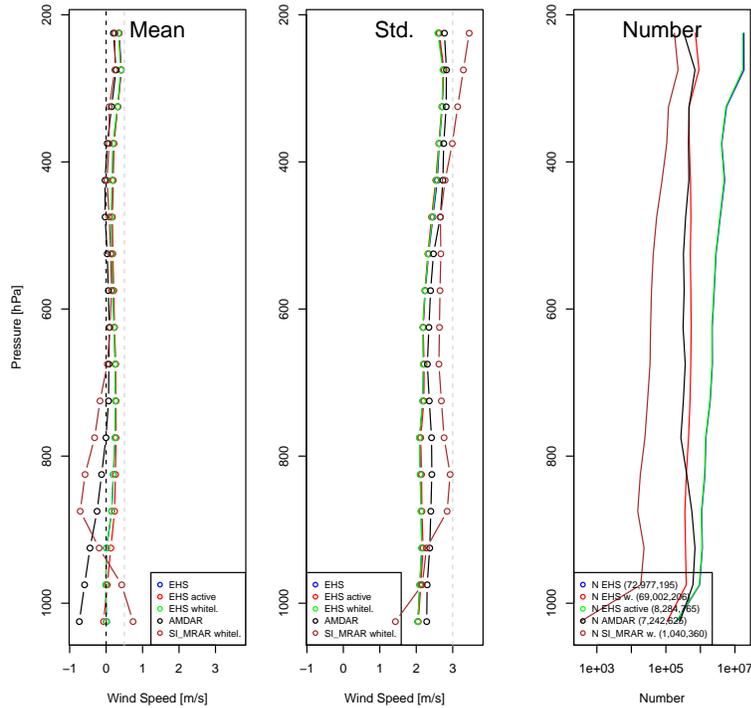
Figure 4: Same as Fig. 3 for wind speed.

this data is only local and in the area mostly outside EHS and AMDAR coverage and that we expect higher forecast errors over the complex terrain of Slovenia which can also contribute to larger wind OMG in the lower atmosphere.

2.2 Whitelist generation

The whitelist generation is based on statistics obtained for individual aircraft. Various thresholds for absolute bias and standard deviation were considered. Some of the obtained results obtained by various thresholds used for blacklisting in Table 1 (requiring 1000 observations per aircraft using 10% subset, this is 73 mio of all EHS observations). It can be seen how blacklisting affects the number of used observations and the number of aircraft.

The final selection is highlighted with blue color in Table 2. The final whitelist is an intersect of individual whitelists per variable, so that only aircraft with acceptable quality in all variables are accepted. The final merged whitelist contains 4355 aircraft identification names.

3 Data thinning

In the operational ALADIN BlendVar at CHMI the thinning distance for aircraft data is set to 25 km. This value was determined in the scope of re-tuning observation error for
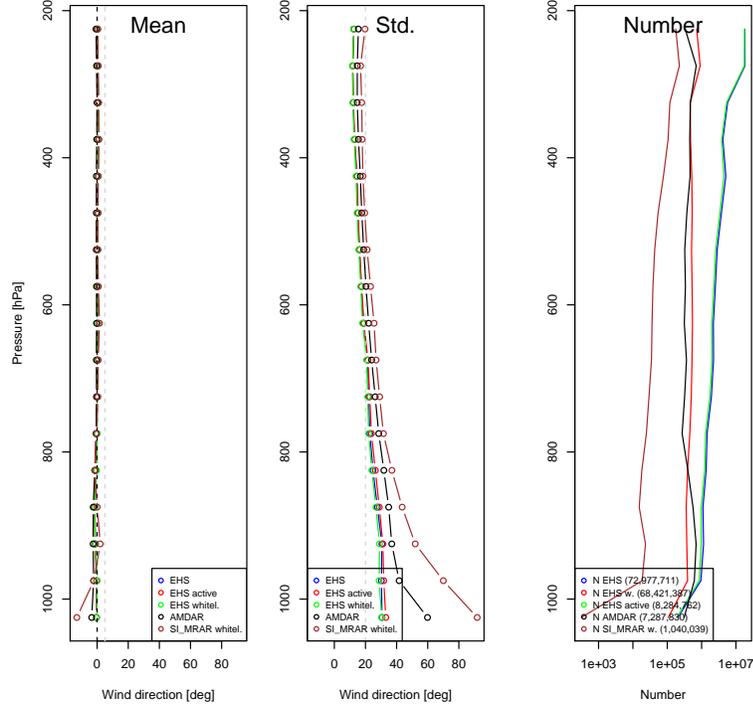
Figure 5: Same as Fig. 3 for wind direction.

Table 1: Whitelist statistics for various thresholds.

var.	gross	mean	sd	N_{data}	$N_{aircraft}$
T	15	0.3	1	53,763,057	3181
T	15	0.5	1.5	70,661,437	4618
T	15	1	2	71,787,439	4728
wsp	50	0.3	2	26,691	16
wsp	50	0.5	3	68,986,593	4038
wsp	50	1	5	72,267,829	4746
v	360	3	20	68,459,623	4494
v	360	5	20	68,480,174	4506
v	360	5	30	72,313,569	4761

The total number of observations per variable is 73,049,218.

aircraft observations in 3D-Var (Benáček, 2016). Observation error multiplication factor (SIGMAO_COEF) was determined and was changed in the operational setup from 0.67 to 2.8.

3.1 Thinning distances

Thinning of aircraft data in ALADIN/ARPEGE is performed separately by the so-called flight paths, i.e. the observations that have the same

- observation type

- observation code type

- observation instrument type

- station id, i.e. the aircraft address or statid

- vertical layer (index of quasi model level).

Due to such selection EHS and AMDAR observations will be treated independently even though the aircraft is in fact the same. This can potentially lead to close and correlated observations. Observations are sorted by thinning boxes of size AIREP_RFIND, and within each box the observations are sorted by the following priority

$$p(i) = 100 * j_b + (9 - n_a) + \frac{1}{100000} * abs(t(i) - t_a(i)) \tag{1}$$

where j_b is the number of thinning box, n_a is the number of active reports (i.e. temperature and 2 wind components give 3 active reports) and $t(i) - t_a(i)$ is time difference from analysis time. Figure 6 shows screening performance for AMDAR and EHS observations from a single aircraft and using 10 km horizontal thinning distance. The active observations after screening are marked by large circles. It can be seen that screening is independent for AMDAR and EHS. We can also observe many nearby EHS observations. The explanation of this lies in the distribution of thinning boxes (vertical and horizontal lines) and in fact that the leftmost observation will always be taken because of the time criterion in Eg. 1. This effectively decreases the thinning. To overcome this issue and select observations lying more towards the center of the thinning box, Eg. 1 can be extended to

$$p(i) = 100 * j_b + (9 - n_a) + (0.5 - d_x(i)) + (0.5 - d_y(i)) + \frac{1}{100000} * abs(t(i) - t_a(i)), \tag{2}$$

where $d_x(i)$ and $d_y(i)$ are relative distances of i-th observation location from the border of the thinning box (computed in thiar.F90 routine). The effect of such modification on data selection in screening is presented on the right panel of Fig. 6.
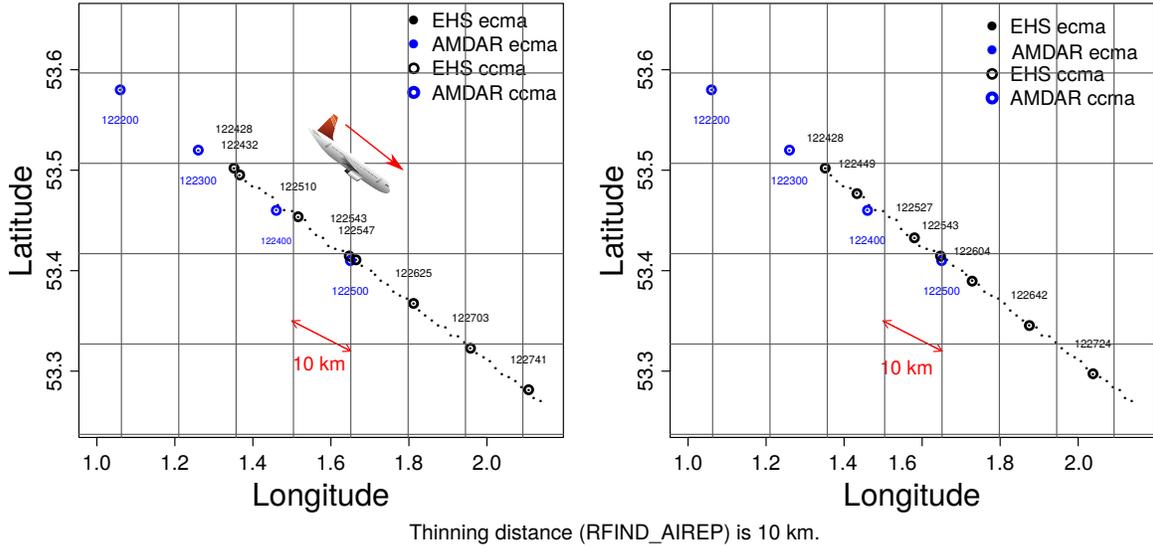
Figure 6: Screening performance for a single aircraft's AMDAR and EHS observation. Left original, right corrected.

3.2 Diagnosing observational error correlation

Given the density of Mode-S EHS observations it is a relevant question how much thinning one can allow not to break the fundamental assumption of non-correlated observations errors in 3D-Var. To diagnose optimal thinning we here apply Desrozier's diagnostics (Desroziers (2005)) and use the same tool as Benáček (2016). For this we set up a non-cycled assimilation experiment (x71, see Table 2) with reduced horizontal and vertical thinning (RFIND_AIREP=5000., RAIREPTHIN=500, RAIREPPCENTTHIN=0.01).

The diagnostics by of Desroziers (2005) is based on OMG and OMA (observation minus analysis) departures. Therefore we can only use active observations. The product of pre- and post-analysis departures is a statistical estimation of observational error covariances

$$\mathbf{R} \approx E[OMA(OMG)^T]. \tag{3}$$

The independent estimation of (parts of) the observation covariance matrix $\mathbf{R}$ in this way is only possible if the true background errors and the true observation errors have sufficiently different correlation structures. The calculation for two groups of observations separated by the same distance d and with locations i,j gives the following error correlation for a distance d:

$$
\begin{aligned}
cor(d) &= \frac{cov(OMA_i, OMG_j)}{\sqrt{var(OMA_i)var(OMA_j)}} \tag{4} \\
&= \frac{E[OMA_i OMG_j^T] - E[OMG_i]E[OMG_j]^T}{\sqrt{(E[OMA_i^2] - E[OMG_j]^2)(E[OMA_i^2] - E[OMG_j]^2)}}. \tag{5}
\end{aligned}
$$

The expectation operator $E[]$ can be approximately evaluated by computing the average of a large sample.

The OMA and OMG pairs were binned in 10 km intervals from 10 to 100 km and correlations were computed according to the Eq. 5. The observations are less than 1 hour apart in time. Also, the pairs were diagnosed separately on the predefined vertical levels of 4 hPa. Only higher levels (traveling altitudes) from 185 to 290 hPa allow the diagnosis of larger horizontal correlation of aircraft data. Figure 7 shows observation error covariances for all aircraft data (EHS, MRAR, AMDAR) as diagnosed in the non-cycled experiment with reduced thinning. It can be seen that horizontal observation correlation exists and it is falling slowly from 0.75 to 0.5 at around 20-30 km (depending on level) and further decreases towards 0.25 at distances from 50 to 90 km. According to Liu and Rabier (2003) the value should be below 0.2 to ensure that is will not compromise the quality of current 3D/Var algorithm.

The values here are substantially larger that those diagnosed in Benáček (2016) for Czech MRAR data. This discrepancy was further investigated. Comparison of diagnostics with reference experiment without EHS data gave approximately the same results. Also effect of increased observation error standard deviation was tested for Czech MRAR data, see Fig. 8. We can conclude that increased observation error standard deviation also causes larger diagnosed error correlations, but this effect is not fully understood and needs further evaluation.

4 Impact on forecast

4.1 Experiments

To evaluate impact on forecast, 54-hour production forecasts are run twice per day. Outputs are saved hourly for up to 24 hours and every 6 hours afterwards. The forecasts are verified against AMDAR and, to some extent, the radiosonde observations. To do this we apply a modified version of the Veral verification package which is used operationally at CHMI.

Table summarizes the considered impact experiments. Experiment x70 uses EHS observations without any modification in the preprocessing nor the whitelisting. Experiment x72 uses the same setup but EHS observations are whitelisted based on the results of non-cycled experiment x71 as described earlier. Experiment x80 does not assimilate any EHS observations and thus serves as a reference for verification. The modified thinning and enlarged thinning distance are applied in x73. Finally, x81 is the same as x72 (whitelisted observations, default preprocessing) except that we only assimilate EHS wind observations.

4.2 Impact of EHS data versus reference

The most visible difference in the forecast scores is between experiments x70, x73, x81 (using Mode-S EHS as is) and the reference experiment x80 (Figs. 9-11). Temperature RMSE and STD are slightly but consistently improved at all forecast ranges, most visibly up to 8 hours. The impact is mostly high/level (850 hPa) and above. Bias is also improved till hour 5. The improvements are even larger for wind speed (0.1 m/s in standard deviation) and also wind direction where bias is more mixed (not shown). From the daily evolution of scores (Fig. 12) at hour 3 of the forecast it can be seen that improvements are occasionally present during the whole simulation time. The profile of relative RMSE improvements against the reference is predominantly negative which indicates consistent improvement (Fig. 13).
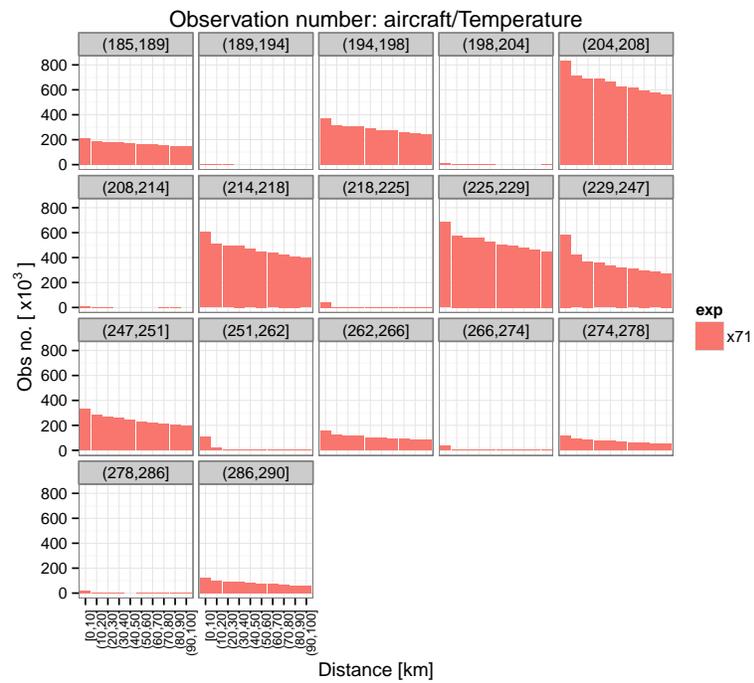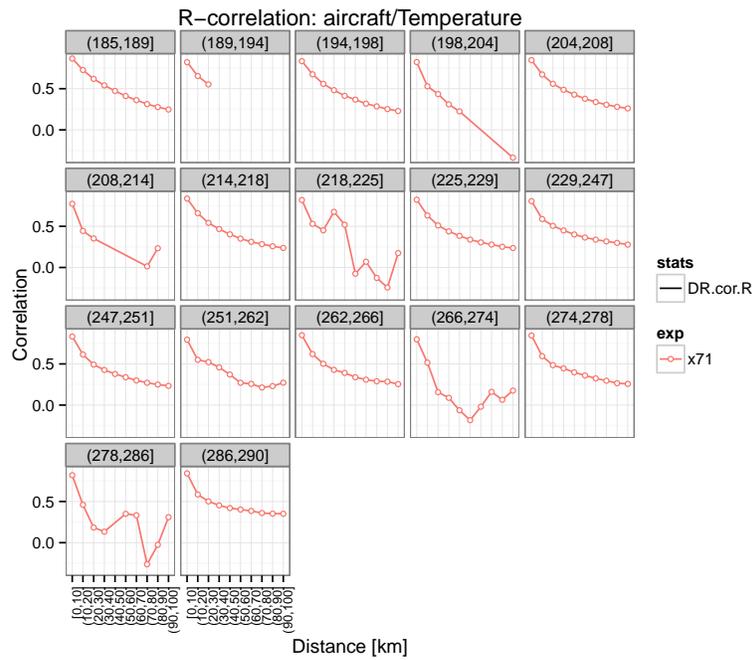
Figure 7: Diagnosed observation error correlation (left) and number pairs for non-cycled experiment with reduced thinning as a function of distance interval.
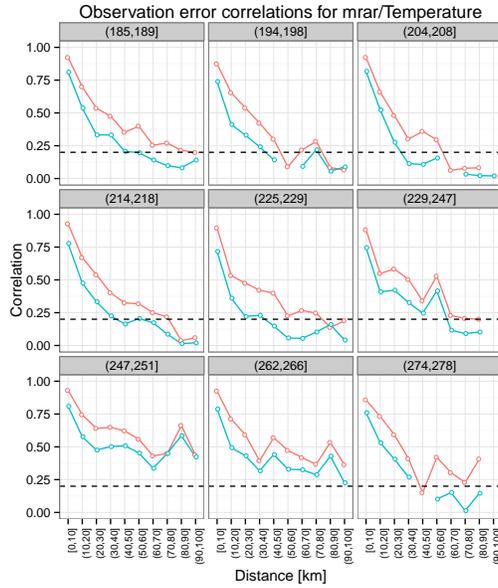
Figure 8: Diagnosed observation error correlation for CZ MRAR data using 0.5 (green) and 2 (red) observation error multiplication factors. Figure by P. Benáček.

Table 2: Experiment list.

| name	period	description	hor.thin.	cycled	assim.EHS
x70	20170111 - 20170210	all Mode-S EHS	25	y	u,v,T
x71	20170111 - 20170210	all Mode-S EHS	5	n	u,v,T
x72	20170111 - 20170222	whitelisted EHS	25	y	u,v,T
x73	20170111 - 20170210	whitelisted EHS, increased thin. distance, adapted thinning	50	y	u,v,T
x80	20170111 - 20170210	reference	25	y	-
x81	20170111 - 20170210	whitelisted EHS, winds only	25	y	u,v
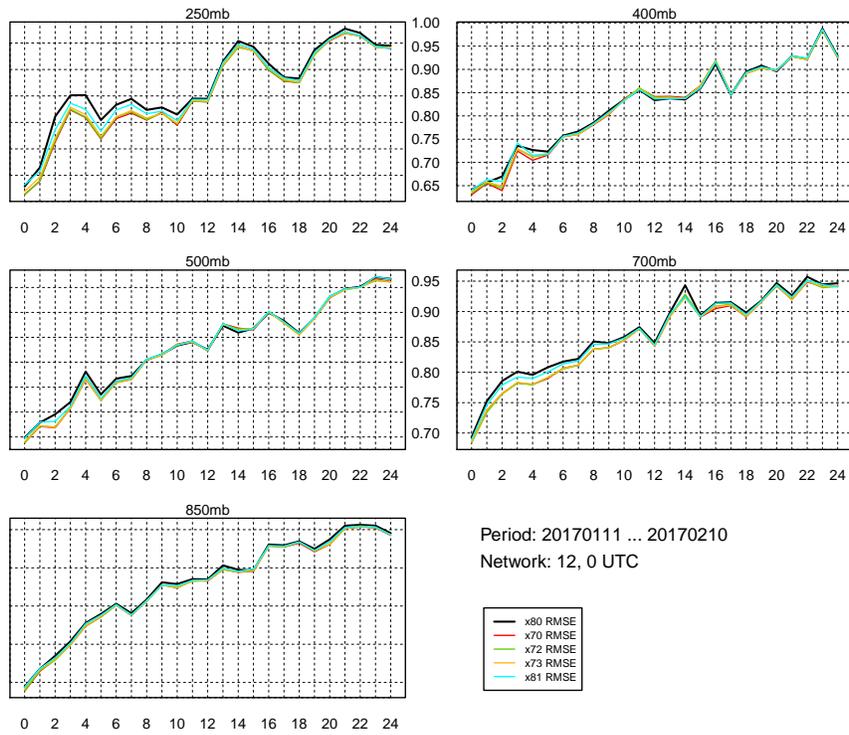
Figure 9: Temperature RMSE comparison of experiments x80, x70 and x72, x73, x81.
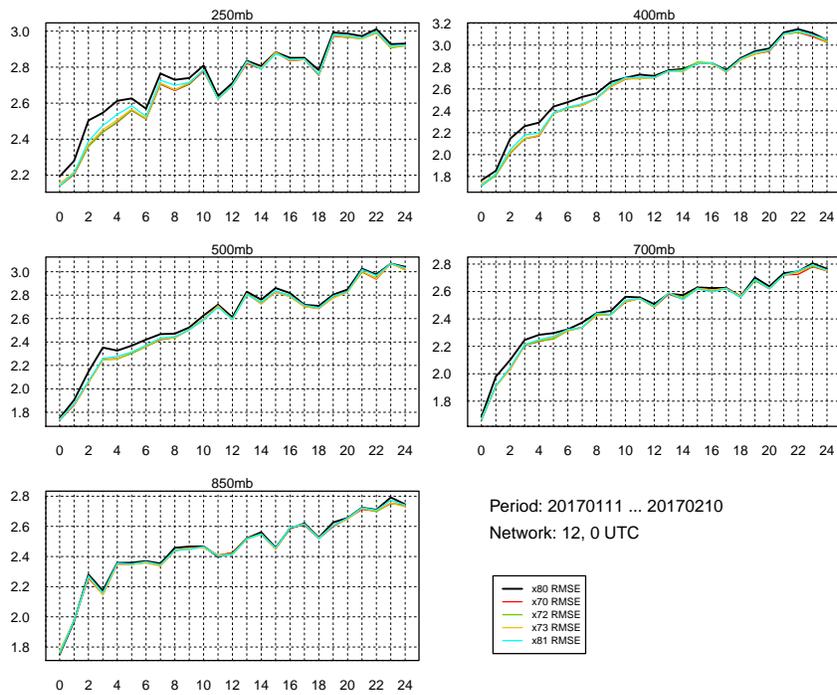

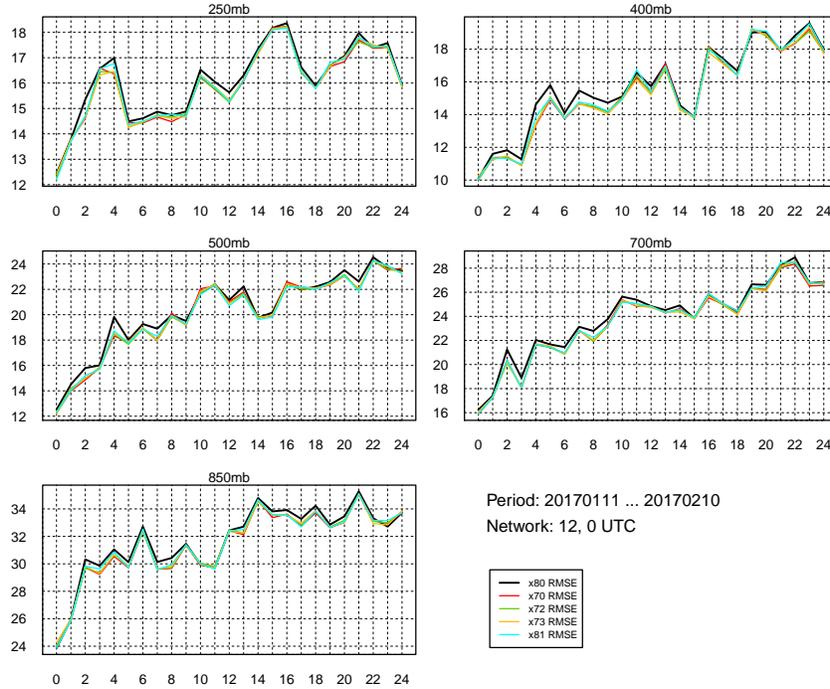Figure 10: Wind speed RMSE comparison of experiments x80, x70 and x72, x73, x81.

Figure 11: Wind direction RMSE comparison of experiments x80, x70 and x72, x73, x81.

The verification was also performed against radiosonde observations at forecast ranges 0, 6 and 12 and the impact here appears to be rather neutral (Figs. 14-16).

Small improvements can be also seen on the precipitation fields in one case. Fig. 17 shows 6 h precipitation accumulation on 14 January 2017 18 UTC. Although both simulations overestimated the precipitation amounts there is less precipitation in x70 in the west-central Czech Republic in the 6-hourly forecast and the precipitation pattern in SW of the country is sharper and more precise and the maximum precipitation extends less northwards. Some improvement in location of the same precipitation band can be observed also in the previous, 12-hourly forecast.

4.3   Impact of whitelisting

The impact of whitelisting as presented in the earlier sections is assessed by a comparison between experiments x70 and x72. The scores show a very minor impact of such pre-selection of EHS data. A very small overall degradation probably stems from assimilating slightly less data in x72.

4.4   Assimilation of EHS wind only

Given the somewhat better quality of wind than temperature, it is interesting to investigate the forecast impact of assimilating winds only (experiment x81). This experiment corresponds to cyan color in Figs. 9 - 11. It can be observed that temperature RMSE is very close to the score of the reference at analysis time and then starts to decrease towards the other
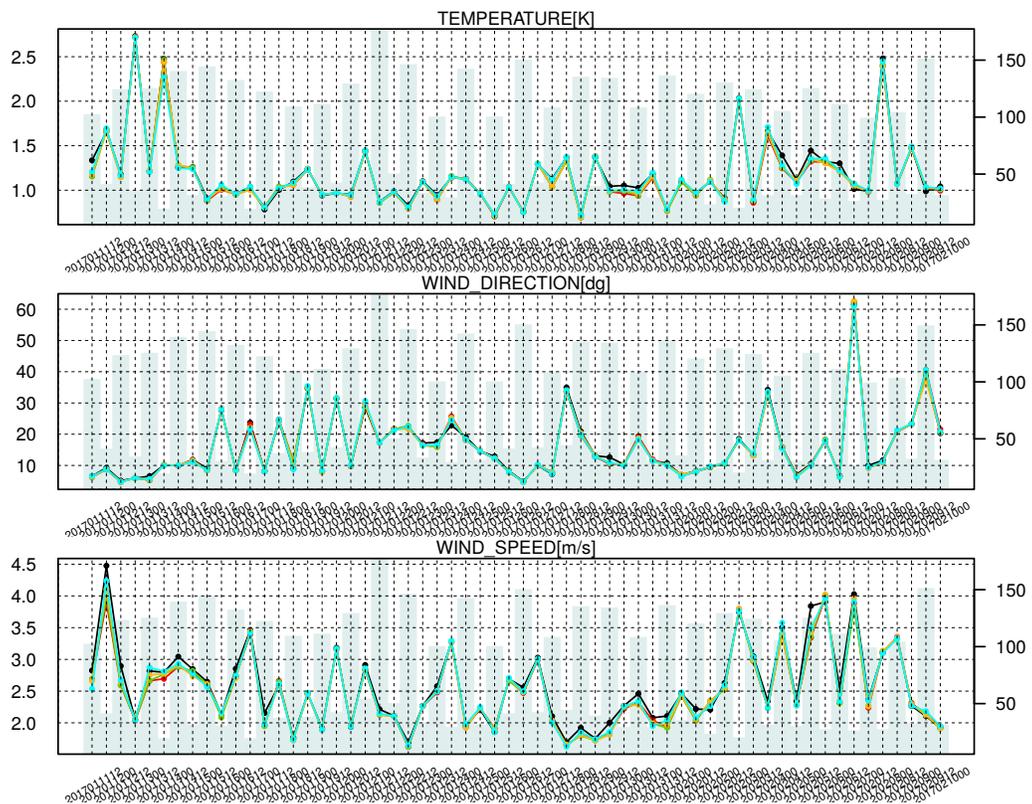
Figure 12: Evolution of forecast scores for all parameters at 500 hPa for experiments x80, x70 and x72, x73, x81.
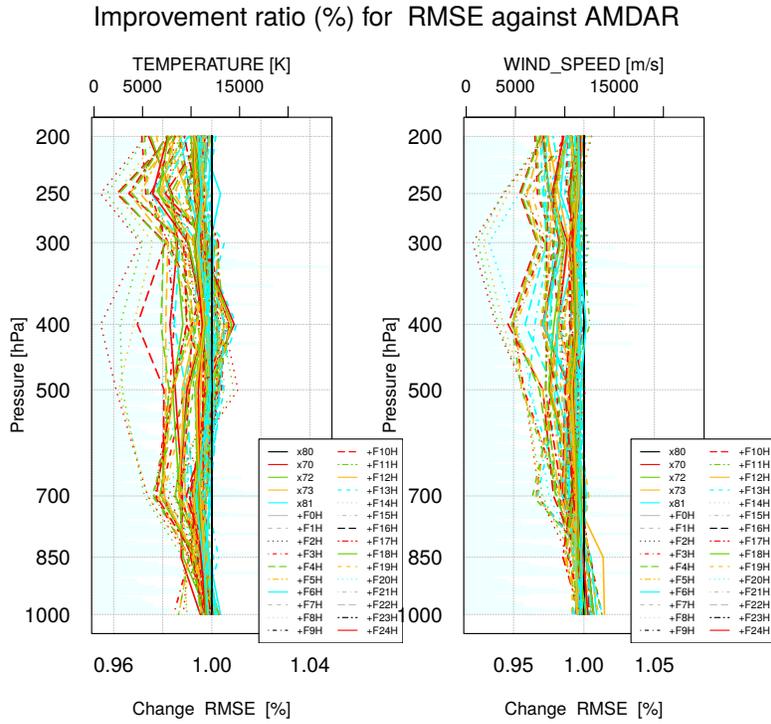
Figure 13: Relative change of forecast RMSE (top) due to assimilated Mode-S EHS.
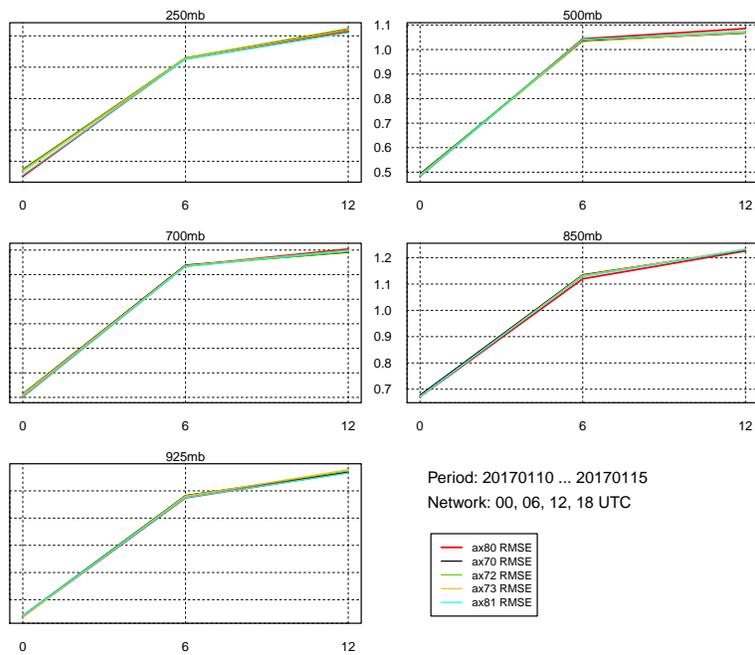

Figure 14: Temperature RMSE comparison of experiments x80, x70 and x72, x73, x81 over a few day verification period against radiosondes.
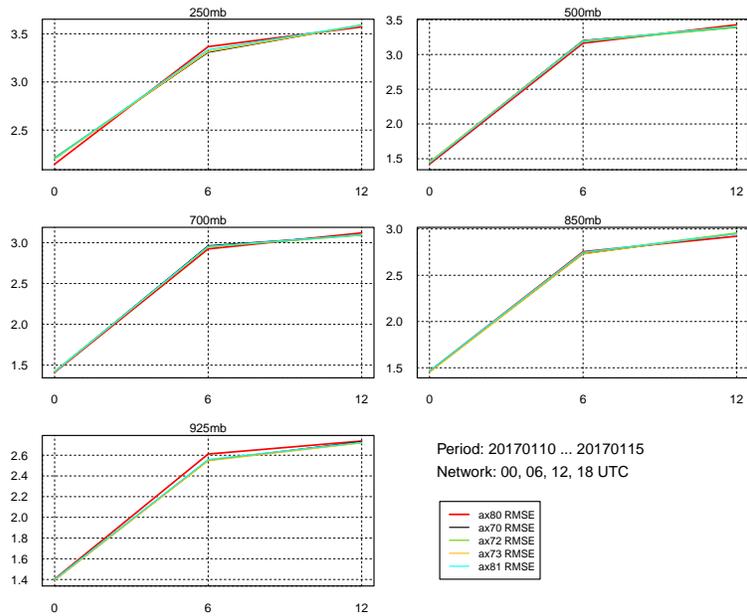
Figure 15: Wind speed RMSE comparison of experiments x80, x70 and x72, x73, x81 over a few day verification period against radiosondes.
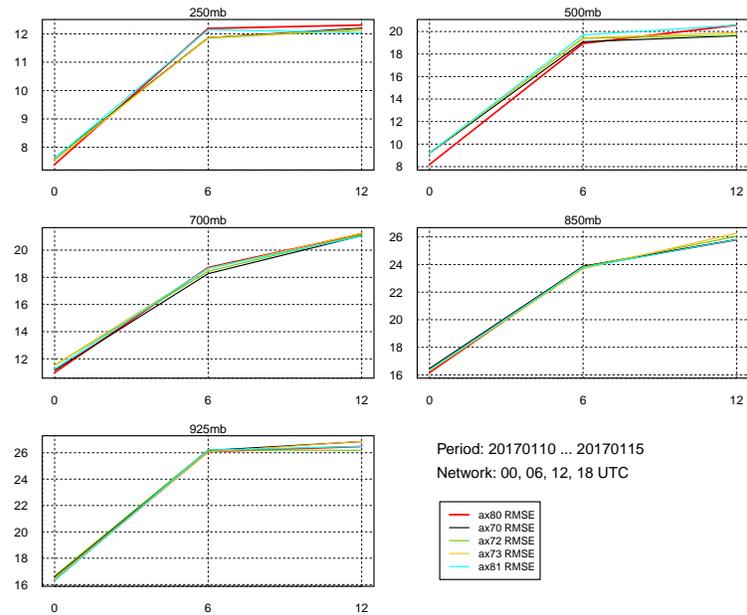

Figure 16: Wind direction RMSE comparison of experiments x80, x70 and x72, x73, x81 over a few day verification period against radiosondes.
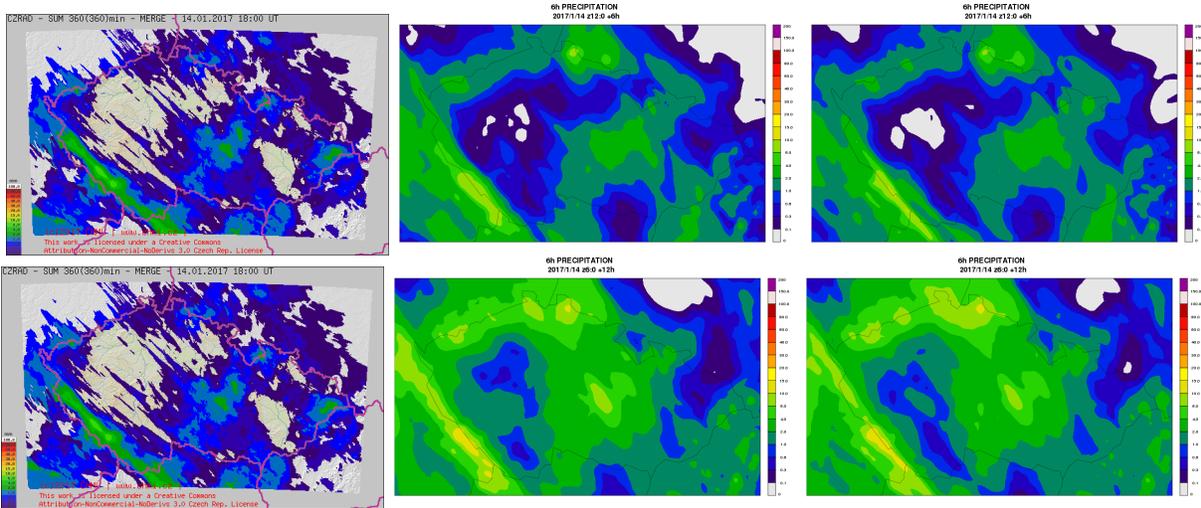
Figure 17: 6h precipitation accumulation [mm] from a reference radar+raingauges analysis (left), reference forecast x80 (middle) and forecast using EHS (x70, right). Shown are the latest 6 hourly forecast (top row) and the forecasts initialized 12 hours before the accumulation time (bottom row).

experiments using EHS at 250 hPa and 500 hPa. On the contrary, wind RMSE is closer to the other experiments using EHS at the analysis time but the RMSE increases faster than for the other experiments during the next forecast hours. The result shows that temperature assimilation also improves wind forecast as the forecast evolves.

4.5 Increased thinning and modified thinning data selection

The impact of tuned thinning can be seen by comparing experiments x73 and x72 in Figs. 9 - 11. It is small but systematically negative, especially for temperature. For experiment x73 we observe the smaller fit to AMDAR of all the experiments which use EHS data. This is reasonable because x73 uses around half less Mode-S EHS data.

5   Conclusions

Our study of Mode-S EHS data was based on a comparison of Mode-S EHS data with Czech operational forecast (ALARO with BlendVar analysis) over the period of 10 months. Five assimilation experiments using EHS data plus the reference with slightly different setups were carried out and two 54-hour production forecast per day for one month period were computed and verified.

A comparison of obs-minus-guess statistic showed good quality with respect to AMDAR. Wind bias and standard deviation of EHS was even better than that of AMDAR, while temperature standard deviation was increased towards lower levels. Bias of temperature was close to zero. The EHS proved to improve both temperature and wind forecast for the whole 12 hour forecast range, as verified against AMDAR data. Verification against radiosondes was neutral. Although EHS seem to have slightly improved precipitation for one of the precipitation cases (subjective assessment), this is not enough to draw firm conclusions on precipitation as the proper objective verification over the complete data set was not done.

The whitelisting approach was minimally reducing the total number of observations. This led to very minor, yet consist degradation of scores. This indicates that also the data outside the thresholds for blacklisting contributed positively to the forecast. We conclude that it is safe to use EHS data as is, without additional prior selection based on individual aircraft statistics. We have to keep in mind that such approach is used at KNMI at the earlier processing stages.

Experiment with denied EHS wind data showed that both temperature and wind contribute to improvements seen in the forecast. In such experiment, the forecast RMSE for temperature decreases with time compared to the reference forecast but stays higher than in the assimilation experiment using both atmospheric variables.

The currently least understood is the result with the reduced thinning as suggested by observation error correlation diagnostics. The impact was slightly negative. However, the optimal distance as diagnosed by Desroziers method seems to depend on the setting of observation error standard deviation, based on comparison with previous results on Mode-S MRAR data and using smaller error standard deviation. The impact might be small in our case because a rather high obs. error multiplication factor is used here (SIGMAO_COEFF). It remains unclear how to tune both SIGMAO_COEFF and thinning distances independently.

References

Benáček P., 2016. Tuning of 3D-Var ALADIN-CZ system for aircraft data assimilation. RC-LACE web page: http://www.rclace.eu/File/Data_Assimilation/2016/rep_Patrik_ModeS.pdf

Liu Z.Q. and Rabier F., 2003. The potential of high-density observations for numerical weather prediction: A study with simulated observations. Quarterly Journal of the Royal Meteorological Society, 2013, 129-594: p 3013-3035.

Desroziers G. et al., 2005. Diagnosis of observation, background and analysis Error statistics

in observation space. Quarterly Journal of the Royal Meteorological Society 131.613 (2005): 3385-3396.

de Haan S. and Stoffelen A., 2012. Assimilation of High-Resolution Mode-S Wind and Temperature Observations in a Regional NWP model, Weather and Forecasting, 2012, 27, 918-937, doi:/10.1175/WAF-D-11-00088.1.

de Haan S., 2011. High-resolution wind and temperature observations from aircraft tracked by Mode-S air traffic control radar, J. Geophys. Res., 2011, 116, D10111, doi:10.1029/2010JD015264.

A Locations of the main scripts on the server kazi1

On yaga server, scripts for processing 10 months of aircraft data are located on /home/maa/mma233/ehs_nl/scripts:

- preparation of sampled OMG departures for EHS data as R database: prepare_rdb.R, sampling.R

- preparation of OMG departures for AMDAR and MRAR as R database: prepare_rdb_AMDAR.R, merge_AMDAR_MRAR.R

- statistics and inter comparison of different aircraft types: qc_monitor_comparison.R

Assimilation experiments:

- x70, x71, x72, x73 are stored on archiv under user mma233

- x80, x81 are stored on archiv under user mma153

Veral verification scores and auxiliary aggregated data: /home/maa/mma233/work/veralR.