

Reviewer 1

General comments:

The paper describes a system that provides wind and temperature data based on aircraft measurements. The European Meteorological Aircraft Derived Data Center (EMADDC) system processes Mode-S Enhanced Surveillance data which is transmitted based on interrogation by ATC radars. Sophisticated methods to minimize errors caused for example by deviations of magnetic heading or truncation errors are introduced. The provision of dense and accurate fields of wind and temperature data in the airspace is of high relevance for aviation meteorology and numerical weather prediction. While the methods to minimize errors and the respective mathematics is quite impressive, there is still space for improving readability of the manuscript. Overall, the paper is a little exhausting to read. One simple means to improve readability could be to revisit the placement of paragraphs. But it would certainly be good to invest some time to improve the writing flow of the text in general. Section 8 is partly not very carefully written. Can you please outline the advantage of EMADDC against AMDAR data? Why do we also need EMADDC? Is it the measurement coverage or frequency? The accuracy of both methods appears similar.

response The advantage of Mode-S EHS is that is available (almost) everywhere aircraft go (in Europe). The costs of receiving data are limited, in most area's the data can be obtained through ATC-partners. The disadvantage is the the temperature is not good for heights below 850hPa. We need a good mix of both Mode-S EHS and AMDAR. The text is adapted likewise.

Specific comments:

The sizes of the labels of several figures need to be increased.

response done

The flow charts in Figs. 1, 2 and 7 are not very nice graphically. Fonts are very small with respect to the space which is not used. The arrows partly overlap words and look like in improvised sketches.

response: New flow charts are made

Table 1: Flight levels are usually described in hundreds of feet. Here, possibly a term like flight altitude may be more appropriate?

response adjusted

Table 1: Does the reported precision correspond to one stand deviation? How are the values for the precision estimated?

response the reported precsion is how many digits are used to describe the parameter value.

Table 1: The global height-keeping performance by ICAO (ICAO, Manual on a 300 m (1000 ft) Vertical Separation Minimum Between FL 290 and FL 410 Inclusive, International Civil Aviation Organization, Doc 9574, AN/934, 3rd edition, 62 pages, 2012) specifies height-keeping errors beyond 90 m (300 ft) in magnitude to less than 0.002 probability. Assuming a Gaussian distribution,

this would yield substantially higher values for one standard deviation than the 25 ft reported here. Maybe for the paper one should state that different error estimates are around and the estimates used here may be connected to substantial uncertainty? Please comment.

response the term reported precissions is ambiguous

Table 2: It would be interesting to know whether these values just indicate errors or whether they also exclude phases of flight (like curved flight or approach phase) that would not allow deriving data of good quality or that are not of interest.

response looking into any correlation between phase of flight and these flags is an interesting idea and might be looked into in the future. These flags are also being refined in our work for the neW EMADDC 3.0 system

Table 2: The exceedance of roll angles of 2.5° in 16

2.5 degrees of roll is relatively small and can even be measured during turbulence. In our opinion, this number is representative. The 2.5 degree limit value is also checked to see if it can be increased in EMADDC 3.0

section 5: The parameters used in the equations should be introduced. It should be checked throughout the paper that all parameters are introduced. Alternatively, a nomenclature could be added.

response done

l. 150: The datum y corresponds to the date? According to eq. (11), y is expressed in fractions of a year? Please introduce the definition of datum.

response done

Fig. 3 is not easy to understand. The text talks about corrections within the year 2023. However, the y-axis extends over 3 years? The legend denotes years? Please add the unit. The legend only shows values in the past which makes sense to me as the values of the declination table of an aircraft may be outdated. However, the lower plot shows also corrections pointing several decades into the future. Such corrections probably don't make much sense. Shouldn't they be disabled or sorted out accordingly?

response done and text has been added

section 5.3: It would be helpful not only to derive mathematically why the heading correction is not sensitive to errors of the numerical wind prediction. Could you please briefly explain in words why this is the case.

response text has been added (see 5.3)

l. 210: Numerical weather prediction models may exhibit not only random error but systematic biases in wind speed for certain regions, seasons or altitudes and combinations of these. Isn't there a risk that the correction of magnetic heading in parts corrects deficiencies of the NWP model?

response we added some words on this in section "Numerical Weather Prediction model Comparison"

l. 218: "wind vector from the model" does this refer to the NWP model? (There are also other passages where the term model is used and it is not fully clear which model this should be)

response added NWP

l. 223: “Care must be taken when model based bias correction are applied because model biases themselves might bias the corrected observation.” How can this be done?

response we are looking into this

Fig. 4: One may argue whether deviations on the order of 0.1 m/s need to be corrected. Or is this done mainly to get better accuracies for temperature?

response correct

Fig. 5: According to the text the error is a variance while in the legend it is a standard deviation.

response error and standard deviation

Caption Fig. 6: “Red line depicts the mean difference between model and observation (solid line) and standard deviation of the difference (dashed line).” This is not easy to understand. I understand that the bias is a solid line and the STDEV is dashed. Why is the colour red stressed?

response this was a typo and is corected

According table 2 data is not used under certain conditions at FLs below 50. Is this because at these low altitudes temperature errors appear increased in Figs. 5 and 6? Can one explain why temperature errors are increasing at low altitudes and why the bias changes so drastically from negative to positive values at these heights?

response the reason is that att low altitudes, aircraft are flying slower, both in Mach as well as airspeed and the truncation of both will introduce larger errors a these heights

Couldn't find a description of Figure 8.

reponse: added a reference

l. 337: “Table 3 also shows the statistics of wind and temperature ...” There is no temperature data in table 3.

response corrected

Table 3 provides wind direction errors in the unit m/s!?

response corrected

l. 350: Again, the mentioned temperature data is not in the table. Caption of table 5 should also mention the NWP data.

response corrected

l. 354: “... has mostly a standard deviation lower ..” add mostly, as the statement is not always true

response added

l. 357: the model resolution does not correspond to its grid size?

response changed simpler wording

Caption of Fig. 9: “Left panel shows the mean differences of temperature ...” add temperature ... “the right panel shows the velocity vector RMS with respect to height” add velocity. Unit of height is missing. The value of 3386 appears not correct.

response The value of ”3386” is correct, it has been checked

l. 368: the curves rather look like a flight level resolution of 25 has been applied than a 25 ft resolution?

response corrected, this was a typo

section 8.3: It should be mentioned that AMDAR is based on the same sensors and the differences between AMDAR and EHS arise from the way the data is processed while of course the difference in location and time also contribute to the deviations.

response words have been added

tables 3 -6, Fig. 9 and text: What is labeled as EHS is probably the Mode-S EHS data processed by EMADDC and not the raw EHS data. To avoid misinterpretations EHS should be replaced by EMADDC wherever appropriate.

response done

l. 360 and 396: "Comparison with radiosonde observations showed good quality with respect to temperature when the observation is above 850hPa." There is no evidence shown of this claim. In table 6 this height range does not occur explicitly and the errors above FL 400 are bigger than those below FL 100.

response you are right we changed the 850hPa to 700hPa and added some words on the increasing standard deviation with height

l. 414: please briefly introduce the meaning of superobbing, as this is not a very common term

response we added some words on this

Reviewer 2

General

This is a useful and interesting manuscript. I want to see the final version include statements to the effect: "Despite the care taken there are still residual errors from the heading correction." and "In the long term it should be better to replace the use of Mode-S derived data with direct reporting of meteorological variables (e.g. via ADS-B)." If the authors don't agree with these statements they should be included as comments from a reviewer.

Response : we added these valuable remarks in the conclusions

From Eyre (2016) and my own experience I take the view that any correction method is imperfect and has residual errors. Learning more about Mode-S, such as occasional errors in interpreting BDS messages or positions, as well as the need for corrections, convinces me that direct meteorological reporting would be better. The fact that EMADDC processing works as well as it does is testament to the hard work and ingenuity applied. As to the question of why the Mode-S and AMDAR wind comparisons to the ECMWF forecast fields are very similar I would suggest that this is partly due to the whitelisting procedure and to the air-speed correction. If similar procedures were applied to AMDAR data I would expect the AMDAR statistics to improve slightly. The residual errors from the heading correction have the advantage that (from most points of view) they are random and hence averaging over 10 or more reports from different aircraft in a small area the errors will largely cancel out. In a thought experiment an aircraft with a residual heading error of 5 degrees (say) travelling from A to B will have

a fairly consistent wind direction error, but as it travels back from B to A the direction error will be opposite. This is assuming that wind speeds are small compared to the air-speed. Some work could be done to firm up these ideas.

response at present research on errors is undertaken and a publication will be prepared soon

Detailed comments

L83 'Contrary for local receivers' - 'However for local receivers'

response Done

L84 'Data received from ATC radars or trackers is also processed ...' - 'also' slightly confusing Is 'tracker' the same as 'local receiver'? 'All data have quality control and filtering applied.' would be clearer

response wording changed

L157 'The correction method uses the assumption that the correction is determined by a geomagnetic reference table for a certain datum (or epoch) ...' Why not check this assumption by talking to Airbus, Boeing or others in the aviation industry?

response we have made several attempt to verify this wirh manufactures to no avail unfortunately. Experience and knowledge within our team lead to this assumption however.

Figure 3. Needs clearer explanation. Are the units in years in both parts of the figure? 'maximum heading correction datum' - could be better expressed/explained NB. It is good to see some information on the heading correction applied - but the tails and the noise reinforce the idea that there must be residual errors.

response we improved on the wording

Figure 4. I would expect the NWP winds to be representative of a larger area and thus (I think) to have fewer values near the extremes of the distribution. Is this what you see? (I am struggling to fully understand the figure.)

response The figure is based on in total 12 hours of data and therefor extremes (when present) are not averaged out

L243 'is of worse accuracy' - 'has worse accuracy'

response done

L256 'dashed blue line' - not dashed!

response done

'8.2 Comparison with Radiosonde Observations' Are the radiosonde data being used at the launch position (as used to be standard) or at the position they have drifted to?

response we used the drift

414 'to apply this most efficient' - 'to apply this most efficiently'

response done

Acknowledgements: Add ECMWF for their forecast data?

response of course!

List of major changes

1. The text has been improved using the reviewers comments and by careful reading itself
2. all flowcharts have been redrawn and look better now (we think)
3. variables used in equations are described
4. Figure captions have been checked and adapted
5. we added some words on "direct measurements" vs "indirect and corrected measurements"