Review of: EMADDC: high quality, quickly available and high-volume wind and temperature observations from aircraft using Mode-S EHS infrastructure, by de Haan, et al.

Summary Statement: Let me begin by saying that I am a strong supporter of the use of Air Borne Observations (ABOs) in all phases of operational forecast, from bench forecasters using the data to improve short-range local forecasts and hazardous weather warnings to integration into NWP systems. Many studies have shown the impacts of ABO data in both applications, especially in otherwise data-poor areas or between conventional radiosonde launch times. Aircraft position and movement information included in Mode-S reports could provide a possible alternative to the direct measures of temperature and wind (and in some cases moisture) provided by more established AMDAR program. A major disadvantage of the AMDAR is the cost of receiving the data via air-to-ground communications in some regions of the globe, while a major disadvantage of relying on Mode-S is that meteorological information must be derived from position and aircraft motion information provided in reports that were originally designed for air traffic management purposes. This paper, along with others by the lead author, takes the position that the risks and possible errors in the deriving larger volumes of meteorological parameters override the costs involved in acquiring more directly observed AMDAR data.

Although I found the section of the paper describing the need to correct for the difference between true north and magnetic north to be thorough and well presented, I found that other parts of the paper, especially those related to determining temperature, to be vague, overly optimistic and not only reproducible, but also possibly incorrect for applications in areas of the atmosphere with significant moisture. Some of these same issues appeared in earlier papers referenced in this submission. Overall, I must recommend that the paper be returned to the authors so that they can make major revisions before reconsideration for publication.

Specific Comments: As I read the paper, I kept thinking that the authors might have submitted an earlier version than intended. At the end of line 66, there is a '(?)' in the end of the sentence. What does that mean? The sentence is also conjecture and probably should be eliminated. Throughout the text, there are many instances where statistics that could be quantified are instead replaced by vague adverbs of adjectives, such as the word 'frequently' in the same line. Similarly, in line 53, the words 'not all' should be quantified. As it stands, it could mean that as few as 1% or well over 50% of radars would not meet one of the 2 conditions described in the sentence. Also, Table columns are incorrectly labelled and variables in some of the equations are not clearly defined. Numerous spelling and grammar errors also need to be corrected throughout the paper.

Lines 56-59: With the large volumes of Mode-S observations available, how much information does the inclusion of questionably encoded reports add to the volume of reports from more reliable transmissions? Please indicate how much these reports might degrade the overall quality of the derived data sets.

Line 60: It would be very useful to list the transmitted parameters that are more important in deriving each of the meteorological parameters early in the paper. E.g., it would be helpful for the reader to know ahead of time which observed parameters affect temperature derivations.

Lines 73-86: This section describes at least 3 different means in which Mach number that are used at EMADDC. Please explain to the end user how they can know which of the three options were used for in deriving meteorological data from each aircraft and how that choice might affect the quality of the reports and how much difference each of the 3 methods makes.

Line 89: Do the Mode-S reports include GPS horizontal position reports as well as altitude? This sentence implies that they do not.

Equation 1: It would seem more logical to identify the dynamic pressure at p_d instead of q_t. Also, the variable in equations 1 and 2 need to be defined in the text.

Section 4.2: Since this is not relevant to Mode-S observations, this section is unnecessary. Also, the equation, if used, should be written so it is solved as T=, not T_i=.

Table 1: This table is incomplete and incorrect in places. Frequency and units are missing for position, even though an accuracy was given in their 2022 paper. The labels of the frequency and reporting accuracy columns are also reversed. Also, although time was listed as a coded parameter in the authors 2022 paper, it is not listed here, nor are the precision of the reported value. This needs to be clarified, since the 2022 paper lists a choice of 2 values (1 s or 1 ms, where 'ms' is undefined). Which is used in your current system? If both are used, what impact does that difference have on derived meteorological variables?

Also, no discussion is presented in this or previous papers about how the onboard reports are 'binned' into their reporting precision intervals. Specifically, were the reports simply truncated was software included to determine if the reports were within +/- ½ of the precision interval on either side of the reported value. This information is essential to determine if biases have been introduced in the data compression process.

Lines 104-109: Nowhere in the paper are the common frequencies used for the various parameters used in deriving meteorological information specified. This is especially hard for a reader to guess since the frequency is listed as a range. In their 2011 paper, the authors indicated that a 15 (or 60) second averaging (or linear fit) of Mach number and air speed was necessary to improve derived temperature calculations. That statement is not repeated in this paper. Has this changed? If so, say so and explain why. If averaging is used as part of the calculations, then only 1 derived parameter should be reported during the entire averaging period to avoid correlated errors between successive corrections and the reporting frequency should be adjusted to reflect that change. This need to be clarified and well documented. Also, please show which of the two parameters (Mach number or air speed) benefited more for using the linear fit smoothing process? Also, and probably most importantly, is the question of whether the corrections applied to both parameters in the linear smoothing process are correlated or uncorrelated in instances where the method improved the derived temperatures. (For reference, investigation of Mode-S wind speed that I have done using a small random sample of data provided ECMWF shows observation-to-observations wind speed changes frequently approaching +/-2 m/s between successive 20 second reports, even after applying a 3 sigma QC filter. This variability could have major consequences on the quality of more instantaneous temperature derivations.)

Lines 115-118: Having looked carefully at a substantial amount of Mode-S derived meteorological reports, I recommend that error bounds of 2 standard deviations be used instead of 3. This more conservative approach is especially justified give the large volume of Mode-S reports and will reduce the data volume by no more than a few %.

Table 2: Please describe how and why these limits were chosen, especially for test 7. Also, if Mach number reporting accuracy is .004, why is .001 used in test 6.

Lines 120-143: This section of the paper concerns me most for numbers of reasons. First, and possibly most significant, is that the fact that the speed of sound (and therefore Mach number) is affected by atmospheric moisture. The 2022 paper explicitly states that the effects of moisture are ignored. Although the effect of moisture is indeed small at upper cruise levels, impacts in the bottom several hundred hPa can be significant, increasing the speed of sound in the moist, less dense environments by over 2 m/s. This can in turn affect temperature derivations by as much as a degree and would result in the derived temperatures being more like virtual temperature than a sensible temperature. This system shortcoming must be recognized. In addition, this large of a change in reported Mach number could shift the transmitted Mach number by one or two of the 0.004 reporting precision increments, which could lead to errors in reported Mach number of up to .008, which could impact temperature calculations even further.

Line 129-132: As with all other corrections applied in this paper, please give a typical magnitude and range of values for these corrections. In this case, how much does the static pressure correction affect both Mach number and subsequent derived temperatures?

Lines 134-143: This paragraph is quite confusing. The first sentence states that NWP is being used to correct temperatures, but no explanation is made of what NWP information is used or how it affects the correction. It implies that corrections are needed for each aircraft individually, but no explanation of the reason for this is documented. A correction formula (formula 5) is then presented based on a new set of static pressures apparently derived from NWP fields. The 2022 paper is then references for more details, but I could not find any there. Instead, the paper describes a method of correcting AMDAR temperature observations, not Mode-S temperature derivations. An explanation of the derivation of the coefficients in (5) is essential for this technique to be reproduced by others, along with examples of the magnitudes of the correction that are applied as a function of altitude. As a side point, one possible explanation of the effectiveness of the technique is that, since the NWP height fields on pressure surfaces are derived using virtual temperatures, this correction could be unknowingly accounting for the effects of water vapor. This needs much further discussion.

Lines 201-205: Although the section on correction for true vs. magnetic north is detailed and seems sound, the statement in lines 201-205 not well documented. No printed reference is given for the 'true air speed bias correction' that EMADDC uses, including no indication of the magnitude of the corrections. It also seems to assume that most of the wind errors are in speed and that wind directions cannot be corrected. Is this true? If so, it should be stated directly. Again, without information about how this correction was formulated, the work cannot be duplicated by others. Finally, if a future physical correction method depends on an already corrected temperature, how might the 2 corrections interact?

Section 9.1-9.2: Line 239 refers to the quality of Mode-S wind data using parameters u, v and wind speed. This list, however, fails to include possibly the best measurement of overall wind observations quality, that being the Vector RMS (VRMS), which accounts for both wind speed and direction errors. Although wind speed fits were similar using both NWP and Radiosondes as comparison standards, approximations of the Vector RMS derived from the u and v fits with radiosondes produce VRMS values closer to 3.25 m/s. Comparisons of reports between AMDAR and radiosonde data over the US show wind speed fits less than 2.0 m/s and VRMS values of about 2.5 m/s throughout the depth of the troposphere. TAMDAR reports were substantially worse, with speed fits ranging from 3-4 m/s and VRMS values of 4.5-5.5 m/s. I recommend that you expand your references to include the US intercomparisons, expand your statistics to include VRMS and make specific reference to the lower quality of the individual Mode-S reports when compared to previous AMDAR evaluations. That said, I believe that if the Mode-S reports were amalgamated over periods of 5-7 minutes (the typical time between AMDAR flight level reports), much of the small-scale noise that I have observed will be removed the statistics would improve substantially. Please make the labeling of column headers in Tables 3-6 consistent, more clearly explain the meaning of 'all data' and 'whitelisted and unique' in the caption. Also, please make the layers in Tables 3 and 4 and in Tables 5 and 6 consistent so that the 2 sets of results can be compared directly.

I was happy to see that the authors used both NWP and radiosondes in their evaluations. As stated earlier, reliable reports taken by individual aircraft during ascent and descent could be extremely useful for operational bench forecasts, especially in land areas without radiosonde coverage and impending hazardous weather.

Finally, although the authors have gone through great lengths to in efforts to derive temperature information from the Mach number of air speed observations available through Mode-S, many fewer temperature derivations were made than the number of wind reports that were made. This reasons for these differences need to be explained (and understood) more clearly in the text, including the methods by which nearly 35% of the derived temperatures were rejected.

With this substantial number of questions remaining to be answered, I can only recommend that the paper be returned to the author for major revision and resubmission. I encourage the author to do so.