We kindly thank the reviewer for his comments and questions.
Below you will find our response to these comments and questions together with questions and comments themselves.
Kind regards,
Siebren de Haan , et al.

line 17 'Upper air observations from aircraft are an important source of information for numerical weather prediction (NWP).' Possibly reference Ingleby et al (2021).

» Done: added reference

line 21,22 'ECMWF has introduced an aircraft and flight phase dependent temperature correction (Cardinali et al., 2004).' The ECMWF bias correction of aircraft temperatures was announced in a short newsletter item by Isaksen et al (2012, see below), an update is given in Ingleby et al (2019). Cardinali et al (2004) did -not- discuss aircraft temperature biases. I think the work at NCEP by Zhu et al (2015) should also be mentioned.

» Done: added references

line 29 'The formal difference is slightly smaller than 0.4K (Painting, 2003).' I'm not sure what this means - clarify or delete. It might be worth mentioning the EUFAR workshop on aircraft temperature measurements (Nov 2020): https://www.eufar.net/shared-subjects/s/3fa8510de42844d6b30 (I particularly remember the presentation by Bob Sable on TAT sensors - it seems the industry preoccupation is with avoiding icing of the sensors in extreme conditions and accuracy was a secondary consideration.)

» Deleted

line 40 'Aircraft sensors' I recommend that the book by Wendisch and Brenguier (2012) is referenced in this section.

» we added text to clarify

line 70ff 'Aircraft temperature measurement' This is key and should probably be expanded slightly - to mention the convertion of kinetic energy to temperature (mainly by adiabatic compression within the TAT probe). Perhaps mention typical differences between Ta and Ti. Section 2.5 of Wendisch and Brenguier is useful - it derives equations like (4). WMO seems to be encouraging use of WMO No. 8 'Guide to Meteorological Instruments and Methods of Observation' rather than Painting (2003).

» Done: used the proposed reference

line 91 '2.5 Numerical weather prediction model data' Either here or earlier the geographical domain being used should be mentioned.

» Done: added a description of the region

line 115,116 'Since generally an atmospheric profile has a temperature lapse rate of -6.5 K/km' 'Since average tropospheric profiles have ...' would be more accurate.

» Done

line 123,124 'the temperature is really biased assuming that the bias not related to the time difference, is independent of the flight phase' perhaps 'there is an additional bias term, which may be independent of the flight phase'

» Done

line 145 What is the typical time difference tau? How much does it vary? Is it linked to aircraft type and/or airline?

»

line 147 'possible' - 'possibly'

» Done

line 161 'Thus, when we have an (estimate) of the mapping f-1 we can correct the temperature measurement.' Either remove the brackets or extend them: '(an estimate of)'. More fundamentally I haven't fully understood this mapping, any extra explantion would be welcome. From Figure 1 I think that larger corrections are needed at lower airspeeds - is this correct (and true of other aircraft)?

» Done: we added some text

line 190,191 'Radiosondes are generally launched at the main hours (00, 06, 12, 18 UTC), as required by WMO, with the majority of launches around 00 and 12 UTC (these timestamps represent the observation at a level of 500 hPa at the whole hour)' 'before the main hours' (often about 45 minutes before, but different NMSs vary). I have heard it said that they should reach 100 hPa at about the main hour. BUFR radiosonde reports have the time of each individual level.

» Done: we added some words

line 214 'The most left panel' - 'The left-most panel'

» Done

line 215,217 'The reason for the difference in bias with the time of day is not understood. Assuming that the AMDAR bias is constant we observe that the radiosonde bias, changes over the day from overestimation at 06 UTC to neutral at 12 UTC and underestimation at 18 UTC to slightly underestimation at 00 UTC.' Radiosondes are mainly available at 00 and 12 UTC as already stated. Apart from low levels the sample at 00 UTC is quite small - because there are fewer flights at night. The proportion of cargo flights at night may well be higher? Given the sampling issues for both aircraft and radiosondes I would advise against suggesting a diurnal cycle in radiosonde biases. Radiosondes have larger uncertainty in the -stratosphere- in sunlight (Dirksen et al, 2014).

» Done: we added some words

line 218 'Figure 4' - interesting variation with aircraft type Looking at the ECMWF bias corrections by aircraft type I see even larger biases for some B787 aircraft (US-AMDAR) and small negative (cool) biases for some Airbus aircraft. Please increase the size of the text labels in figure 4 to improve clarity.

» Done

line 225 'the second is an accuracy related correction.' perhaps 'the second comes from the interconnected nature of aircraft measurements: there is a Mach number correction to the temperature and a temperature correction to the Mach number and it appears that avionics systems do not iterate to convergence.'

» Done

line 236,237 'The Mode-S EHS information can be applied to correct the AMDAR temperature bias, for those air spaces where Mode-S EHS information is available.' This is not a long-term solution. The meteorological community needs to persuade the aviation industry to improve their avionics/measurements.

» Done

line 258 'Painting, J. D.: WMO AMDAR Reference Manual, WMO-no.958, WMO, Geneva, http://www.wmo.int, 2003.' http://www.wmo.int no longer exists and WMO regard this document as superseded, see https://community.wmo.int/activity-areas/aircraft-based-observations/resources/manua and-guides