Response to Referee #2 comments

The authors would like to thank the referee for their detailed comments, which have led to significant improvements in the overall clarity of the paper.

General response to Referee's comments:

The aim of this paper is to establish a methodology, from the limited data available from sites with a high launch frequency, to see if a data correction factor could be established at these sites to guide launch schedules. This represents the first step in developing a general tool for calculating temporal correction factors for any ground based monitoring site.

The aim is not to physically explain or quantify the reasons behind the correction factors derived at each site.

The above has been added to the introduction to clarify the scope of the work contained in the paper.

The following text has been added to the conclusions section:

"Having established that the method presented in this paper is a viable one for estimating temporal variability it should be recognised that these results only directly apply to the radiosonde launch sites from which the datasets have been obtained. In order to generate appropriate correction factors for other sites the method will require further development, using additional data sources or model results for each site."

1) Referee's comment:

My first major concern is the lack of physical interpretation of the results. Here are some thoughts on physical interpretation that aren't discussed: Lindenberg and Manus have radically difference climates, with SGP somewhere in between. How applicable are the results from one location to the others? The manuscript especially needs some discussion of the relative amplitudes of the diurnal cycles over Manus, over northern Europe and over Oklahoma. In a similar vein, the vertical variations in the differences shown in Figure 2 suggest some sort of standing wave structure. Does this have implications for later results? Also, what about the annual cycle? If the diurnal cycle is small but the annual cycle is large, then it could explain some of the difference in the figures of seasonal mean.

Author's response:

The clarification to the scope of the paper by the additional text in the introduction and conclusion sections addresses this comment. The correction factor determined is specific to the site for which it was derived. In order to generate appropriate correction factors for other sites the method will require further development.

2) Referee's comment:

My second major concern is the applicability of the proposed methodology. Basically, the estimation method is sound, but the reason for that will not be obvious to non-specialists. The paper is motivated by a laudable need to estimate the errors from temporal mismatch. However, this analysis is applicable to particular sampling timing: observations taken at systematically different local times, as happens when a sun synchronous satellite is compared to synoptic radiosondes. If the measurements

have random time differences (say, dedicated sondes against satellites, or, GPS RO against sondes or other satellites), then the temporal mismatch adds a random error. This study addresses only systematic mismatch errors. The manuscript needs to be much clearer about the nature of the proposed correction, and its applicability be made much more obvious. This is touched on in on p. 8341 around line 25, but it should be clearly stated that the method used here requires a systematic time difference between sondes and satellite to be valid. On the plus side, this analysis is appropriate for comparison between sun-synchronous satellites and all operational sondes launched synoptically– not just those intended for satellite comparisons. Another strong selling point is that the time difference between sun-synchronous satellites and sondes at synoptic times 0Z and 12Z varies monotonically with longitude. Quantifying this effect is a strong justification for this work – but it is not mentioned in the manuscript.

Author's response:

The authors disagree with this comment on the introduction of a random error linked to a 'random' temporal mismatch. The method of calculating the correction factor is based on specific radiosonde launch times at specific locations and provides a measure of the temporal temperature gradient during particular periods. If the temporal mismatch is known between the radiosonde launch and the comparison result at that specific location, time and season then the correction factor can be calculated for that specific mismatch. The correction is therefore not tied to specific radiosonde launch times or mismatches.

3) Referee's comment:

That said, this applicability should be weigh against the small-proposed correction factors of order 0.1 K. This is significantly smaller than the typical errors from all satellite instruments except GPS RO. This needs to be part of the discussion.

Author's response:

The referee's comment is incorrect as the correction factor is determined per hour of temporal mismatch. If the temporal mismatch is large, say, 6 hours, then the correction to the recorded temperature can be large, i.e. > 1 K. In addition, as highlighted in Figure 6, the uncertainty in the correction factor can be significantly higher than 0.1 K for a small number of samples.

4) Referee's comment:

The discussion of the methodology is incomplete. The description of Figure 3 simply states the correction factors "were calculated". Similarly, for Figures 4 and 5 the meaning of the term "calculations repeated" is unclear. The correction needs to be discussed in much more detail, maybe including an explicit definition with an equation. As for Figures 4 and 5, sondes are launched every 12 hours yet the calculations are shown for every six hours. Presumably these are time-corrected, but the discussion is missing. More explanation of the basic calculation methodology is needed before this manuscript can be published.

Author's response:

The last part of the comment has been address first. All of the data analysis on the Lindenberg and Southern Great Planes results was from radiosondes launched 6 hours apart.

It is agreed that the methodology requires a more complete description. Additional text has been added to Section 3 along with equations to clarify calculation process. This text is given below:

Radiosonde temperature readings are amalgamated into altitude bins 500m high, labelled as the centre of each bin, i.e. 0 to 500m labelled as 250m. The temperatures in each altitude bin are averaged to provide a mean temperature, *T*, for that specific altitude. The rate of change in temperature between single launches 3 hours, 6 hours and 12 hours apart, at each altitude, were calculated according to Eqn 1. The mean rate of change in temperature between each launch separation and altitude, $\frac{dT}{dt_n}$, were then calculated according to Eqn 2.

$$\frac{dT}{dt_n} = \frac{T_n - T_0}{t_n - t_0} \qquad \text{Eqn 1}$$

Where *n* = 3, 6, or 12 hours separation between launch time.

$$\frac{\overline{dT}}{dt_n} = \frac{\sum \frac{dT}{dt_n}}{i} \qquad \text{Eqn 2}$$

Where *i* = the number of launch pairs

The mean rates of change in temperature ($\frac{\overline{dT}}{dt_3}$, $\frac{\overline{dT}}{dt_6}$ and $\frac{\overline{dT}}{dt_{12}}$) were used to define temperature change profiles over the day at different altitudes and are shown in Figure 2.

In order to quantify the difference between the different launch schedules it was assumed that 8 launches per day provided the best available measure of the changing state of the atmosphere. The mean hourly rates of change in temperature from these launches were therefore considered to be the base set. The difference in temperature change rates, $\Delta\frac{\overline{dT}}{dt_n}$, (K / hour) between the base set and a single launch, 2 launches a day and 4 launches a day were calculated according to Eqn 3. The results of which can be seen in Figure 3.

$$\Delta\frac{\overline{dT}}{dt_n} = \frac{\sum ABS\left(\frac{\overline{dT}}{dt_n} - \frac{\overline{dT}}{dt_3}\right)}{8} \qquad \text{Eqn 3}$$

Where *n* = 6 or 12. For single launches, $\Delta\frac{\overline{dT}}{dt_n}$ was taken as the mean of $\frac{\overline{dT}}{dt_3}$.

5) Referee's comment:

Finally, most of the results in Table 2 and Figure 6 are simple rescaling by dividing the first set of standard deviations by the square root of 30, and by 10. This does not justify their inclusion in the paper, and a simple discussion would suffice to summarize this result. Similarly, the conclusions twice state that 4 sondes launches are needed to resolve the diurnal cycle. Since the Shannon sampling theorem states the same thing (more or less), this is not an important result.

Author's response:

The referee is correct in stating that the standard deviation of the mean reduces by root n as the number of measurements, n, increases. However we feel that Figure 6 gives a good visual indication of the uncertainty compared with the magnitude of the correction factor, and so have retained the figure.

Information on the statistical significance of the differences between a single launch and 4 launches per day has been added using the text below:

"In Figure 3, the 4 launches per day data set is statistically different from the single launch data set at all altitudes except 3250m, with a confidence level of 1σ (68%). At the 2σ (95%) level, 3 altitudes (9250, 12250 & 15250m) are statistically different."

Additional comments

6) Referee's comment:

Abstract first sentence. A preposition is missing, probably "of".

Author's response: Implemented, 1st sentence now reads:

"Radiosondes provide one of the primary sources of upper atmosphere temperature data for numerical weather prediction, the assessment of long-term trends in atmospheric temperature, the study of atmospheric processes and provide a source of the intercomparison data for with other temperature sensors e.g. satellites."

7) Referee's comments:

Abstract body. Say where the data were acquired, mention limitations to fixed time differences.

Abstract line 12. Another missing preposition.

Author's comments:

Agree that the data sets should in named in the abstract and that the correction factor is between two given launch times. Due to several comments on the abstract by Referee #1 and Referee #2 the full text of the revised abstract is given below:

"Radiosondes provide one of the primary sources of upper atmosphere temperature data for numerical weather prediction, the assessment of long-term trends in atmospheric temperature, the study of atmospheric processes and the intercomparison with other temperature sensors e.g. satellites. When intercomparing different temperature profiles it is important to include the effect of temporal mismatch between the measurements. To help quantify this uncertainty the atmospheric temperature variation through the day needs to be assessed, so that a correction and uncertainty for time difference can be calculated. Temperature data from an intensive radiosonde campaign at Manus Island in Papua New Guinea were analysed to calculate the hourly rate of change in temperature at different altitudes and provide recommendations and correction factors for different launch schedules. Using these results, three additional longer term data sets were analysed (Lindenberg 1999 to 2008, Lindenberg 2009 to 2012 and Southern Great Plains 2006 to 2012) to assess the diurnal variability of temperature as a function of altitude, time of day and season of the year. This provides the appropriate estimation of temperature differences for given temporal separation and the uncertainty associated with them. A general observation was that 10 or more repeat measurements would be required to get a standard error of the mean of less than 0.1 K per hour of temporal mismatch."

8) Referee's comment:

p. 8340, line 19. "Upper atmosphere" should read "upper troposphere and stratosphere".

Author's response: implemented:

"Radiosondes provide one of the primary sources of upper troposphere and stratosphere temperature data…"

9) Referee's comment:

p. 8341, line 14. Randal is a misspelling.

Author's response: Implemented: "Randal" replaced by "Randel".

10) Referee's comment:

p. 8342, line 11. "Diurnal cycle" is more appropriate than current wording.

Author's response implemented:

1st sentence of Section 2 changed to: "To help quantify the difference between radiosonde and satellite measurements the diurnal atmospheric temperature variation needs to be assessed,…".

11) Referee's comments:

p. 8342, line 22. State explicitly how the difference is calculated (later time minus earlier time). See comment below about Figures 4 and 5.

The sentence on top of p. 8344 reads "The correction factor... were calculated".

Author's response:

The additional text and equations added in point 4 address these comments.

12) Referee's comment:

p. 8344, line 13. Four sondes are sufficient to resolve fully resolve a 24 hour cycle. The semidiurnal component of the variability will alias onto the mean.

Author's response:

The difference between 4 launches per day and a single launch has shown to be statistically different in point 5. Due to the lack of sites that launch more than 4 radiosondes per day over a long time period, the data sets from Lindenberg and Southern Great Planes were assessed. If data becomes available in the future for a site with a more frequent launch frequency this assumption could be further validated.

13) Referee's comment:

p. 8345, line 5. Where are the changes in radiosonde type and analysis procedures that could explain the difference?

Author's response:

The data analysis has been performed on publically available data sets. As noted by the referee and commented on in Section 3.2, changes in radiosonde type and post processing routines have not been taken into account in this work, but could be addressed in future work.

14) Referee's comment:

p. 8346. What sort of correction should be applied (presumably bias), and is the included "uncertainty" random?

Author's response:

Agree that the application of the correction factor should be clarified.

The correction factor should be subtracted from the radiosonde measurement. The following sentence has been added to the last paragraph of Section 3.2.

"This correction should be subtracted from the radiosonde measurement to adjust for the temporal mismatch."

The calculated uncertainty is random as it is based on the standard error of the mean change in temperature. Insert the word "random" in front of uncertainty in two instances in the main body text.

15) Referee's comment:

p. 8346, line 10. See comments above about four samples per day resolving the diurnal cycle. Eight will resolve the diurnal and semidiurnal components.

Author's response:

Agreed, but there is virtually no data from radiosondes launched 8 times a day. Our work has shown that 4 launches per day are acceptable but 2 launches per day are too few.

16) Referee's comment:

Figure 4 and 5. The dash in the figures could be interpreted as 00:00 minus 06:00. Also, the caption should define the error bars as is done in the text.

Author's response:

Agreed that the nomenclature is confusing, so we have replaced all occurrences in the main body, Figures and Supplementary data with start time to end time, e.g. "00:00 – 06:00" replaced by "00:00 to 06:00".