Reponses to the comments by Referee 1 are identified below. Note that given that these are 'technical corrections' in the Co-Editors opinion (also see below), we have not gone into great detail of the corrections, but just identified the replacement or additional text where appropriate. Responses to reveiwers comments in bold. Note that we have also added further acknowledgements and corrected some citation errors relative to the last version.

Co-Editor Decision: Publish subject to technical corrections (13 Dec 2016) by Peter Haynes
Comments to the Author:
Both referees recommend that the current version of the paper be published 'as is'. Therefore I am pleased to accept the paper for publication in ACP.

Referee 1 notes a number of further points of clarification that might be made -- some of which seem sensible to me. Therefore please can you consider addressing these prior to providing a final version of the paper. (I am regarding these as possible 'technical corrections' -- the next version of the paper that you provide will simply proceed to publication.)

This paper is a revised submission of the article by Friedrich et al. that deals with comparisons between reanalyzed winds and wind observations gathered during Google Loon long-duration balloon flights. The paper also compares balloon trajectories with those computed with reanalyzed winds. I re-iterate my initial statement on the previous version that this study uses an unique dataset to provide very helpful information on reanalysis accuracies, which are widely used to study transport in the atmosphere. I have furthermore found that the authors have thoughtfully addressed most of my comments on the previous version, as well as those of the other reviewer. I therefore recommend publication of this article. I have nevertheless listed a few minor points that the authors may want to address before publication.

Minor points

- p7, l146 (end of introduction): I still find that the transition between the end of the introduction,

which discusses previous studies, and Section 2 is a somewhat abrupt. I would encourage the authors to either include a short outline of their studies, or a transition sentence.

We have now added the following at the end of the introduction:

The remainder of this paper documents the Loon observations 5 **(Section 2.1), introduces the methodology used in our analysis and specifically details the trajectory model used**

(Section 2.2). Comparison of the Loon zonal and meridional wind speeds with reanalysis products is then detailed (Section 3.1) and the Loon flight paths are used to examine the accuracy of trajectories derived from the reanalyses in Section 3.2.

- p7, l160: remove comma after "while"

Sentence changed to:

While super-pressure balloons typically move along isopycnic (constant density) surfaces during the rare occasions of altitude control this is no longer the case.

- p11,l272 to p12, l283: One likely reason that could explain the differences between this study and Podglajen et al. (2014) regarding reanalysis accuracies in the tropics is that Podglajen et al. (2014) deal with deep tropical balloon flights (within 10◦ of the equator), while this study considers observations mostly southward of 15◦S. It is expected from simple balance argument that mass information provided by spaceborne instruments provides less and less constraints on the wind field as one gets closer to the equator. This is for instance illustrated in Baker et al. (2014) (their Figure 2 notably), which shows that the largest wind errors between models are actually located in the deep tropics (and above oceans).

Added the following sentence:

The fact that Podglajen et al. (2014) also examine a narrower latitude band (within 10^o of the equator) may also be important.

- p13, l317-324: Another possible reason to explain this difference in trajectory separation is that the balloon flights considered in Hertzog et al. (2004) took place in the stratospheric polar vortex. The separation between the real and simulated balloons in this study was therefore somehow limited by the polar vortex size.

After examination of Figure 1 in Hertzog et al. (2004) we are not sure that this point is significant and have therefore not added any text.

- p16, l388: I agree with the authors' statements on my previous comment. Yet, I find that some of the words used around here are inducing some confusion. For instance, the authors use "SZA bias" while they are referring to differences between the temperature of the lifting gas (i.e. in the balloon envelop) and that of the ambient air. On the next sentence, they carry on with "the solar heating on the lift gas temperature is much more significant than the usual solar bias", i.e. the one which takes place when one is measuring the ambient air temperature. I would suggest to make a distinction between:

1. a real measurement bias associated with the daytime radiative heating of temperature sensors (used either to measure ambient air or lift gas temperatures). This bias is mostly dependent on the sensor size, coating and ventilation, and is quite unlikely to explain the 30-K difference between the lift gas temperature measurement and the real ambient air temperature,
2. a physical issue, which is that the daytime lift gas temperature is indeed warmer (by a few tens of degrees) than the ambient air temperature. This issue arises as the balloon envelop absorbs in the UV-visible, and thus conductively heats the lift gas.

My impression is therefore that using Google Loon temperature measurements to infer air temperature is mostly spoiled by this second item than by the sensor bias itself, which was what was corrected for in the previous studies mentioned in the article. I would thus rather use “temperature difference” or anything that clearly identify this contrast. I furthermore notice that this second item likely also explains the nighttime lift-gas temperature that are sometimes significantly colder than the ambient air. Here, it is the envelop absorption in the infrared which plays a role: if the balloon flies over high, optically-thick clouds the envelop cools and cools down the lift gas with respect to the ambient air.

We have rewritten the section to replace SZA bias with differences. Updated text below:

There are several difficulties associated with the Loon temperature data. As previously stated, the data result from measurements of the lift gas temperature and not of the ambient air, resulting in strong solar zenith angle (SZA) dependent differences between the lift gas temperature and the ambient air temperature. These may result from the combination of the daytime radiative heating of temperature sensors and, we speculate, the balloon envelop absorbing in the UV-visible range. Additionally, although we are not aware of the specific instruments used, it seems that the thermometer used has a high uncertainty and is intended as a diagnostic instrument rather than for scientific data collection. An example of balloon–reanalysis temperature differences is show in Fig. 8. The temperature differences between the lift-gas and ambient air can be corrected through the use of a correction function, as is commonly done to adjust for temperature measurement biases arising due to radiative heating of the temperature sensors (Hertzog et al., 2004, 5 2006; Knudsen et al., 2006), but it should be noted that the impact of solar heating on the lift gas temperature is much more significant than the usual solar bias, up to +30 K as opposed to the typical ∼ 1.5 K. The temperature differences can be modelled as:

$$T_{diff} = \alpha + \begin{cases} \beta(1 - e^{(\theta-95)/\lambda_0}) + \gamma e^{-(\theta-90)^2/\lambda_1} & \theta \leq 90 \\ \beta(1 - e^{(\theta-95)/\lambda_0}) + \gamma e^{-(\theta-90)^2/\lambda_2} & 90 < \theta \leq 95 \\ \gamma e^{-(\theta-90)^2/\lambda_2} & 95 < \theta \leq 150 \\ \gamma e^{-(\theta-90)^2/\lambda_2} + \delta \cdot (\theta - 150) & 150 < \theta \end{cases}$$

where α, β, γ, δ, λ_0 , λ_1 and λ_2 are fit coefficients determined from a linear least squares regression. After removing some flights with anomalous observations (unreasonably large differences, questionable GPS or pressure data), we use temperature data from every second flight to fit the correction function, and then apply this correction to the remaining flights. The fitted parameters are provided in Table 4, and Figure 9 shows the CFSv2 temperature differences with and without the correction applied. Application of the correction functions reduces the mean Loon-reanalyses temperature differences to a few degrees, significantly improving the utility of the Loon temperature measurements, however the standard deviation and the shorter term, day–to–day differences are still much greater than observed in other studies.

Ignoring the differences between lift-gas and ambient temperatures by focusing only on the nighttime measurements, we still find a standard deviations of $\sim$ 6 K while other balloon studies typically have biases and standard deviations less than 2 K. Additionally the nighttime measurements show interesting behaviour with common consistent night-long differences of up to ±10 K. Considering the upper bound on the thermometer uncertainty provided by the Loon team, the significant difference which is much greater than those usually dealt with using correction functions, and the unusually inaccurate night-time temperatures, leads us to conclude that currently the quality of the Loon temperature data means it is of little value in assessing the quality of the reanalyses.