This is an excellent paper that describes empirical correction of Aeolus wind bias based on temperature gradients across the primary mirror. The work is important because correction of bias is an important consideration for assimilation of data into numerical forecast models. Two methods are investigated: one based on comparisons of measurements with the ECMWF model, and one derived from measurements of the velocity from ground hits of the transmitted laser pulse. The paper is well-organized and provides details on the correction methodology as well as performance of the correction methods described. Although the results are unique to Aeolus, and therefore are likely of somewhat limited impact for other instruments, the analysis showing the impacts of temperature gradients across the mirror and the conclusion that empirical corrections can be successfully applied are potentially important for addressing unanticipated problems in that crop up in future missions.

Although I think the paper could be published as is, there are a few places in the text where a bit more detail and explanation might be useful to the reader. I leave the decision on whether to request these changes to the discretion of the editor.

Specific (minor) comments on the manuscript:

Line 207: It isn't clear to me how the 86 km averaging of the Rayleigh channel is taken into account when comparing the AUX_MET data with the Level 2B results. The text implies that the nearest neighbor from the model is compared to the L2B data, but the discussion seems unclear to me on issues such as 1) Are the O-B statistics comparing an 86 Km average with a single point from a 9 km grid-spaced data set, and 2) Is the level 2B HLOS measurement placed at the center of the 86 Km swath? Perhaps I'm missing something here, but it seems that some clarification on the details of the comparison would be useful here.

Line 218: The authors should perhaps provide some evidence for the statement "O-B values are averaged over all range gates which is justified by the lack of altitude dependency". One can make a case that the physical effect that creates the temperature gradients won't change with altitude, but it isn't clear whether the statement is based on that assumption or that a comparison was used to make the case for the lack of altitude
dependency.

Line 255: If the bias structure is strongly dependent on the atmospheric scene, that would appear to limit the effectiveness as the scene changes from day to day. I assume that the effects are a function of the time scale of the changes in cloudiness versus the temperature response of the mirror, but perhaps a bit more discussion here could be useful.

Line 268: This is the same issue as noted above. The time scale of the OLR changes would seem to be important if the results from the prior day are being used to correct the bias for a given measurement period.

Line 359: The meaning of the sentence "Note that for the reprocessing data from the same time period is used to derive the fit coefficient." isn't clear to me. Perhaps I missed an earlier reference to reprocessing.

Line 438: In the sentence "Note that the constant offset of about 3 m/s between O-B and ZWC values is due to the different calibration procedure between L1B and L2B winds and is not considered to be a problem for the bias correction since this offset could be corrected in the data processing", reference is made to the different calibration procedures. A reference of a bit of explanation would be useful here.

Line 443: Regression theory is not my specialty but some metric or reference for why 659 samples is not sufficient to use the original model might be useful here.

Line 462: It seems the "and" before "without M1 correction" in the caption for Figure 11 could be eliminated.

Line 477: Use of the dash to indicate the temperature range causes confusion when followed by a negative number in "0.3°C – -0.1°C". Perhaps another way to articulate the range could be employed.

Line 523: It isn't clear to me why it would not have been possible to observe the increase in random error without the bias correction. Perhaps a sentence of explanation here would be useful.