Verification of parameterizations for clear sky downwelling longwave irradiance by Pace et al. uses data collected from Thule, Greenland, to evaluate parameterizations for estimating the downwelling longwave flux at the surface based on screen height meteorological measurements. I have not seen an intercomparison quite like this and I think the results are of interest and publishable in *AMT*. I have a few minor comments for the authors to consider before this manuscript is published.

Dear Dr Cox, we really appreciate your useful comments that focus on and clarify some aspects of the paper. In the following the answers to individual comments:

- A major revision that expands this analysis to include the other YOPP supersites (https://www.polarprediction.net/key-yopp-activities/yoppsitemip/the-yopp-arctic-and-antarctic-supersites/) would make the conclusions more broadly interpretable. However, I respect the scope that the authors are setting and if they do not wish to expand the analysis, I think it would be beneficial to provide some additional text contextualizing the Thule area. For example, northwest Greenland is within a small sub-region of the Arctic where the atmosphere is generally drier (lower IWV) (e.g., Cox et al. 2012) and indeed in the vicinity of 1 cm IWV, there are spectral effects (Cox et al. 2015) that I suspect could impact derivation of coefficients for the parameterizations (more challenging still if one were to interpret Thule data as representative of higher elevations over the ice sheet). Clouds at Eureka, Canada, also in this dry region, are higher, colder, and thinner (e.g., Shupe et al. 2011) compared to much of the rest of the Arctic. Thus, Thule may not be an ideal analogue for either the ice sheet or for the Arctic as a whole. That does not to devalue the results presented here, but is relevant context for readers to understand.

We agree that extending this work to other sites could be useful to people involved in surface radiation budget research in the Arctic, both in the context of the YOPP and in a wider context, given that there are only few detailed works for estimating the DLI in this region. On the other hand, extending the analysis to other sites is not automatic as it would require an in-depth check of the different databases, also thinking about the methodology for determining clear skies, which goes beyond the scope of this work.
We also agree that more details to contextualize the characteristics of Pituffik area (formerly known as Thule) should be provided. We believe it is important to characterize at the best the site in terms of meteorological conditions when carrying out this type of study, so that a possible reader interested in using the results of the paper can do so with knowledge of the facts and it is precisely for this reason that the distributions of the values of e_s, IWV and T_s were shown in the manuscript, implicitly indicating the limits of applicability of the results.

Following your suggestion, a description of the characteristics of THAAO has been added at line 83.
The THAAO is located on South Mountain, at 220 m a.s.l., near the Pituffik Space Base (formerly known as Thule Air Base), along the north-western coast of Greenland at about 3 km from the sea and 11 km from the Greenland ice sheet (GrIS). Therefore, the THAAO environment is typical of the northern coastal area of Greenland, i.e., influenced by both the GrIS which generates strong katabatic winds, and by the sea, especially in summer when open waters prevail over sea ice. Pituffik is also located in a region, which includes the area northwest of Greenland and the Ellesmere Island, characterized by an atmosphere particularly dry (Cox et al. 2012), with higher, colder and thinner clouds with respect to what is found in other areas of the Arctic (Shupe et al. 2011).

- Long and Turner (2008) is essential reading on the topic presented here and should be referenced and considered for this study as well as future work (as indicated at L432-438).

We agree with the reviewer that it was a lack not to cite the work of Long and Turner (2008). Unlike the other parameterizations used in this work, the one presented by Long and Turner (2008) uses variable coefficients not only as a function of physical parameters, but also includes the effort to explicitly represent the daily variability of the parameters in the studied sites. This parameterization therefore requires a more specific approach which involves the optimization of the so-called "Lapse Rate Coefficient" (a coefficient that depends on the lapse rate originally defined by Brutsaert, 1976) by evaluating its variations before sunset, during the night and after dawn, and interpolating in time the obtained results. Given the different approach with those presented in our manuscript, we chose not to test this parameterization which requires more study and introduces site-specific daily variability. Similarly, we decided not to

include the methodology presented by Dürr and Philipona (2004), who optimize the Lapse Rate Coefficient by taking into account the periodic annual and daily variability of the studied sites.
On the other hand, the parameterization presented by Jin et al. (2006) is also based on the concept of the Lapse Rate Coefficient, and it is expressed as a function of the T_s only, derived by an analysis of more than 700 radiosoundings launched from the Arctic station of Resolute Bay. This parameterization has been used and discussed extensively in this work.

Accordingly, the two papers by Long and Turner (2008) and Dürr and Philipona (2004) have been included in the manuscript, providing a brief description of their approach in line 53.

The parameterizations in Dürr and Philipona (2004) and Long and Turner (2008) differ from those considered in this work because they use explicit dependences on the annual and daily variability of the observed atmospheric parameters and DLI at the measurement site and therefore require specific analyses. Both the works improve the parameterization of atmospheric emissivity presented by Brutsaert (1975) by refining the estimation of the so-called Lapse Rate Coefficient. Dürr and Philipona (2004) approximated the diurnal and annual cycle of the considered sites using a periodical function, while Long and Turner (2008) analyzed separately the daytime and nighttime behavior of the Lapse Rate Coefficient interpolating the daily results during sunset and sunrise; they also applied this method to the Arctic site of North Slope in Alaska, finding differences within ± 4 W/m^2 between the measured and observed DLI values in 68% of cases.

- Could you explain in more detail (more quantitatively) the accuracy of your clear-sky detection method (L67-69)? Could you clarify if it is necessary to capture all instances of clear-sky or (I think) only to capture a large sample of confidently detected clear-sky? Could you clarify the sensitivity of the vulnerability in this method to assigning "clear-sky" to cases with high, cold, optically-thin (e.g. cirrus) clouds?

The choice to use the pyrometer instead of the pyrgeometer itself to determine clear sky conditions, presents the advantage of a larger sensitivity to the presence of thin clouds, but the disadvantage of a reduced portion of sky detected at the zenith; this is the main factor that makes it difficult to associate a quantitative accuracy to the developed clear-sky detection method.
Simulations of the pyrometer zenith brightness temperature (IBT) and of the DLI have been carried out using the MODTRAN5.3 radiative transfer model, to evaluate the uncertainty associated with the presence of thin cirrus clouds. The aim was not only to verify the sensitivity of the pyrometer to the presence of cirrus, but also to quantitatively determine the influence of these clouds on the DLI.

The main characteristics of the simulations are summarized in the following table.

Atmopheric profile	Subarctic winter	Subarctic summer
IWV	0.3 cm	1.2 cm and 1.5 cm
Acloud base altitude	8 km	8 km
cloud base temperature	220.6 K	239.2 K
Geometrical depth	1 km	1 km
Cloud type	cirrus	cirrus

The MODTRAN internal cirrus model, called cirrus standard model, is based on ice particles with 64 µm effective radius. Overcast conditions are assumed. The cirrus optical thickness values have been of respectively 0.03, 0.1, 0.3, 1, 2, 3, 5. The values of 0.03, 0.3 and 3 has been chosen considering the pioneering work of Sassen and Cho (1992), who defined these thresholds to define sub-visible, thin and opaque cirrus clouds. A winter and two summer cases has been simulated, whereby 0.5 cm and 1.2 cm are the average seasonal values, and 1.5 cm is used to assess the sensitivity of IBT and DLI to larger column water vapor below the cloud.

The increase in the value of DLI and IBT as a function of the cirrus optical thickness is presented in Figure1a for both the winter and summer case. The labels close to the lines show the increase in the value of DLI and IBT compared to that of the clear sky simulation.

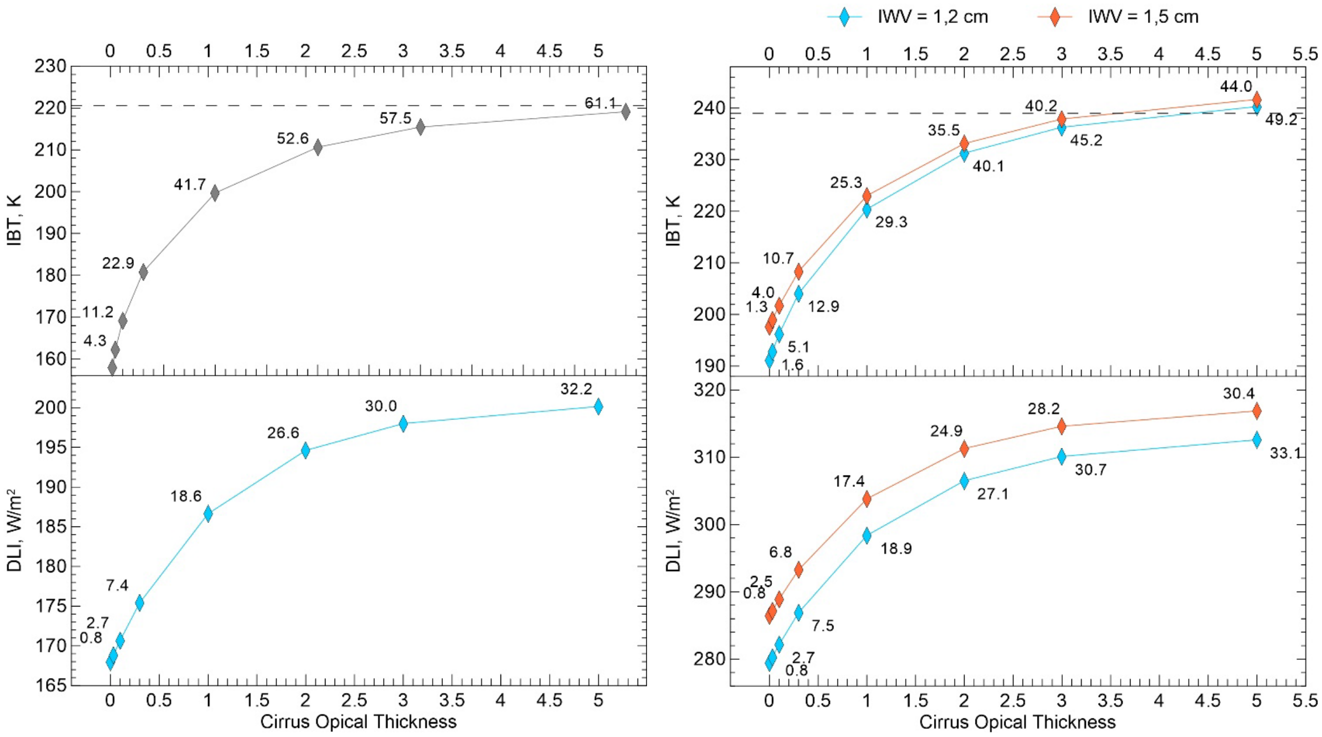

Figure 1a. The plots show the DLI (in the bottom panel) and the IBT (in the top panel) increase as the optical thickness of the cirrus cloud increases, for the winter (left) and the summer (right) case respectively. The labels show the increase in the value of DLI and IBT compared to that of the simulation in the absence of cirrus. A horizontal dotted line highlights the temperature of the cloud base in the winter and summer cases.

The clear-sky IBT for the winter case is 158 K, which is below the pyrometer's calibration range, i.e. down to 173 K, but within its measurement range, i.e. down to 123 K. Even the presence of cirrus with optical thicknesses of 0.03 and 0.1 determines an increase in the pyrometer signal of 4.3 and 11.3 K, respectively equal to an increase in IBT of 2.7% and 7.1%, respectively, compared to cloud-free conditions. Both these values are clearly visible compared to the background signal and within the pyrometer measurement uncertainty that, considering the temperature at the base of the cirrus, is approximately ±1.6/2.0 K. On the other hand, the increase in DLI with respect to its cloud-free value (i.e. 168 W/m^2) is just 0.8 (0.5%) and 2.7 (1.6%) W/m^2, that is particularly small also taking in to account the uncertainty on DLI measurements that are estimated to be ±5 W/m^2.

Similar results are found for the summer simulations. In this case, two simulations were performed taking into account different values of IWV. For IWV equal to 1.2 cm, even considering cirrus optical thickness of just of 0.03 and 0.1 the IBT increases, compared to that with clear sky (i.e. 191.1K), by 1.6 (0.86 %) and 5.0 (2.66%) K, respectively. It should be kept in mind (see paragraph *2 Site and measurement* of the manuscript) that, by decreasing the temperature difference between the cloud base and the pyrometer, the IBT associate uncertainly (±1.3 K) decreases. The same cirrus cloud determine an increase in the clear skies value of DLI, i.e. 279.5 W/m^2, respectively of 0.8 (0.3%) and 2.7 (0.97%) W/m^2 for 0.03 and 0.1 optical depth. Further increasing in the IWV does not substantially change the results, although the decrease of the IBT sensitivity to this cirrus is larger than that of DLI.

In summary, the simulations highlight the larger sensitivity to the presence of thin clouds of the pyrometer compared to the pyrgeometer, especially in the polar environment characterized by low IWV and therefore larger transparency in the atmospheric window around 10 µm.
However, our method for defining clear skies is based more on the variability of the signal than on its intensity.
These simulations highlight how the better sensitivity of the pyrometer induces a larger variability of the signal which is therefore more suitable than that of the pyrgeometer to be used to define clear sky conditions, also because it is much less influenced by the IWV changes (see Figure 2a).
Furthermore, it must be considered that zenith measurements are generally more sensitive to the 2D spatial variability of the cloud than hemispheric measurements. It is therefore unlikely that the algorithm can indicate a *zenith clear sky case* condition corresponding to the presence of a cloud that significantly influences the DLI.

Considering the results of the simulations and the visual inspection of the data (e.g. Figure 2a and discussion in the following), using a conservative approach, the methodology correctly evaluates the presence of cirrus clouds with an optical thickness larger than approximately 0.07-0.1 that, depending by the atmospheric profile (mostly IWV), and the physical (cloud base height, geometric thickness) and microphysical characteristics (ice content, size, shape...) of cirrus should determine an increase in the IBT value of no less than 7/15 K.

Considering the interest that both reviewers have shown in the clear sky algorithm and to better answer their questions, we present and discuss an example of how the algorithm operates to recognize clear sky cases, or perhaps it would be better to say cases where the presence of clouds does not appreciably influence the DLI.
Figure 2a shows six days of IBT, DLI, IWV and T_s measurements collected in February 2018; cases recognized as clear sky are shown as green points.

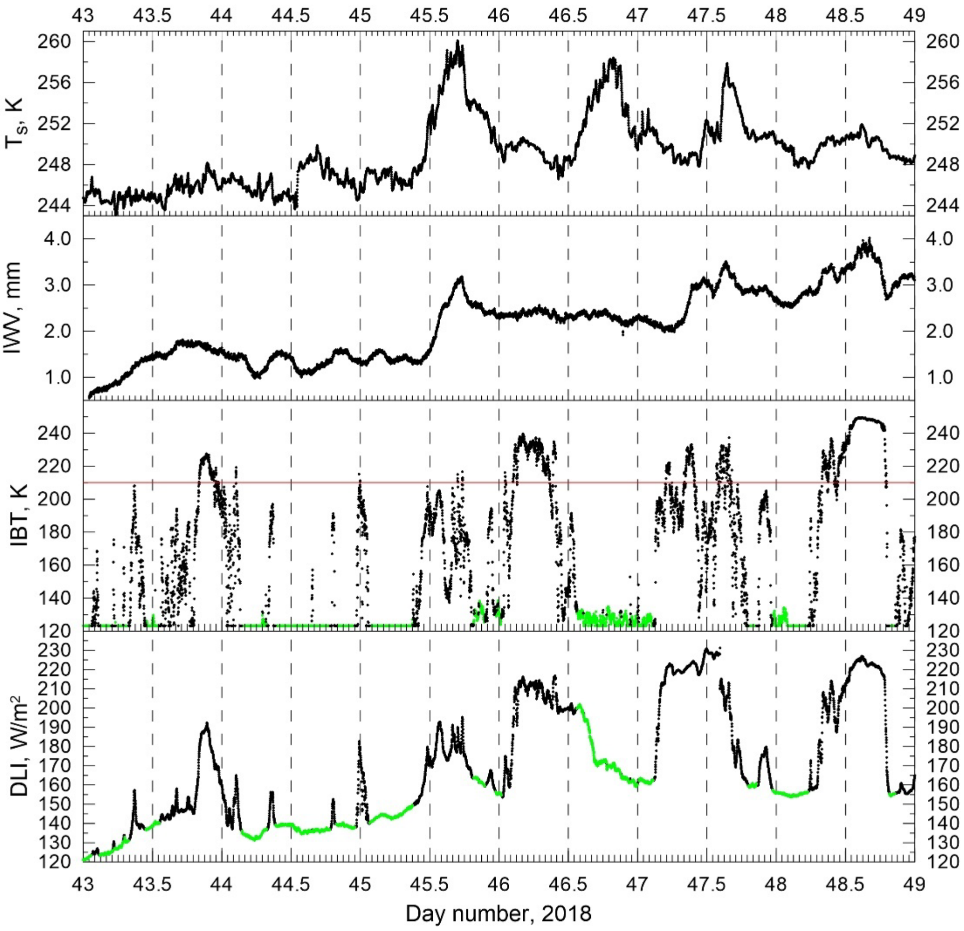

Figure 2a. From bottom to the top: time series of DLI, IBT, IWV and Ts for the period from day number 43 to 49 of 2018, i.e. 12-18 February. As expected, the behavior of DLI is strongly dependent by the variability of IWV and Ts, which do not always show similar patterns.

The algorithm proves to have enough sensitivity to detect very thin clouds that only slightly increase DLI, but determine a IBT increase of approximately 15 K or less (e.g., see the IBT on the beginning of day 43). Although the algorithm proves to work well in relation to the purpose of this analysis, we have chosen to present this case to highlight what has already been mentioned in article lines 167-175, i.e. the importance of visual control of the clear sky.
In the time interval from 46.55 to 46.7 the IBT does not highlight any clouds, while the DLI shows a decrease suggesting residual coverage of the sky. Although this hypothesis cannot be discarded, observation of the sky images shown in Figure 3a suggests another explanation.
Just as happened on the dome of the sky imager between 6 and 18, it is probable that some frost condensed on the dome of the pyrgeometer and then slowly sublimated. This can occasionally occur even if the pyrgeometer is ventilated. In addition, the BSRN quality tests on DLI may not detect such effect. It should be remembered here that the surface of the window of the pyrometer is ventilated with an air flow coming from inside the observatory and is therefore less subject to these phenomena.
These phenomena and the presence of snowfall were the main obstacle to the correct and automatic functioning of the algorithm. Thanks to a visual analysis of the entire dataset their impact is considered negligible.

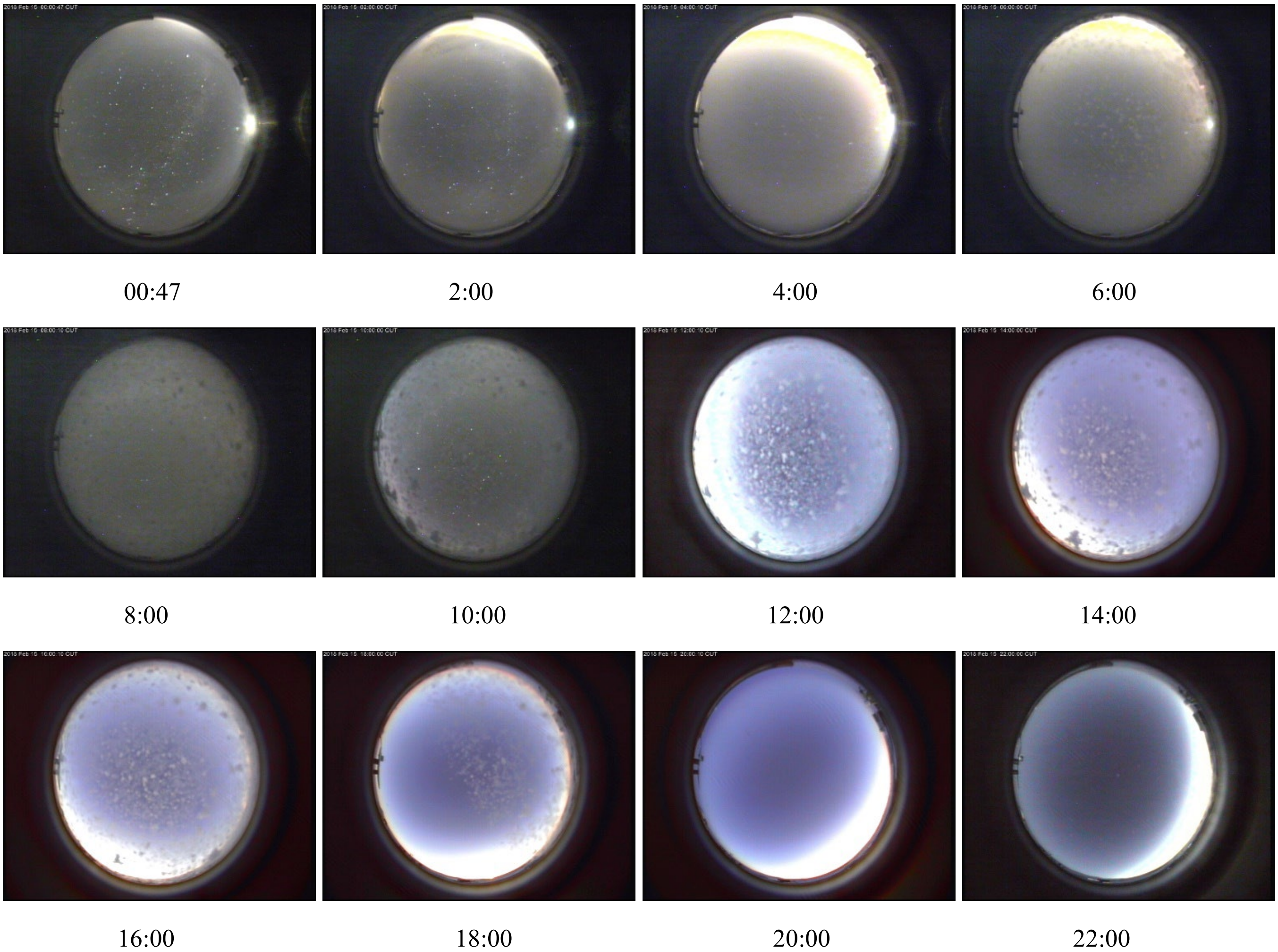

Figure 3a. Bi-hourly sky images of day 46 of 2018 (15 February 2018).

Also taking into account these effects it is much more complex to quantitatively demonstrate that the passage from a *zenith clear sky case* condition to a *clear sky*, i.e. hemispheric clear sky, does not include the residual presence of clouds. The choice to consider a relatively long series of *zenith clear sky cases* to define a *clear sky* condition (please remember that our method identifies a clear sky condition, e.g., at 12:00 if on the 61 IBT measurements ranging from 11:30 to 12.30 at least 45 zenith clear sky cases occur) is based both on the visual analysis of the dataset and the conclusion of the interesting work by Kassianov et al. (2004) that states ”for a relatively short averaging time (15 min), the zenith-pointing observations with a narrow FOV (lidar/radar) can greatly (more than 100%) overestimate/underestimate the cloud fraction”.

Summarizing the above to directly answer the reviewer's question:
Could you clarify if it is necessary to capture all instances of clear-sky or (I think) only to capture a large sample of confidently detected clear-sky?
The choice to define a clear sky measurement by evaluating 61 *zenith clear sky measurements* was maybe a little bit conservative, but (also taking into account the large dataset used) it was preferred to remove some clear sky data rather than include dubious data.

Reference

Sassen, K. and Cho , B.Y. :Subvisual-thin cirrus lidar data set for satellite verification and climatological research, *J. Appl. Meterorol*., 31, 1275–1285, 1992.

Kassianov, E., C. N. Long, and M. Ovtchinnikov (2004), Cloud sky cover versus cloud fraction: Whole-sky simulations and observations, *J. Appl. Meteorol.*, **44**, 86–98.

Some editorial comments:

- L37: Ohmura et al. (2001) is a good reference here as well.

The reference has been added to the sentence.

- Figure 1: I assume the straight lines in 3 & 4 (just before 2018.75) and the bottom panel (between 2017.75 and 2018.00) are artifacts of the plotting technique? Some other short periods in year 2 bottom panel as well. Can these be removed so we can properly visualize data gaps?

Done

- The word “which” is used frequently in places where the correct word is “that” (e.g., L16, 24, 25 but then throughout the text).

The correction has been implemented throughout the text.