RC2: ['Comment on amt-2023-181'](), Claudia Di Biagio, 07 Oct 2023

First of all I would like to sincerely apologize for the delay of my revision.

The paper by Pace et al. investigates the performances of different parameterizations to estimate the clear–sky downward longwave irradiance (DLI) based on comparison against ground–based observations at the high latitude station of THAAO, Greenland. A set of 17 empirical formulations for the DLI are tested. The authors use two full year observations (2017 and 2018) with the aim of: 1/ evaluate original formulas against 2017 data; 2/ use 2017 data to provide optimized coefficients for the THAAO conditions; 3/ test the optimized coefficients against 2018 data. Observations at THAAO used in the present analysis include IBT (infrared zenith sky brightness temperature) and DLI, as obtained from pyrgeometer and pyrometer instruments, and meteorological parameters. The pyrometer data are used to derive cloud–free periods based on an original algorithm, while the pyrgeometer data are used as direct comparison against calculated DLI with the 17 different formulations.

The paper is well written, clear and logically organized. The dataset, the procedure, the figures, and the results are clearly presented and discussed. The final scope of this work, that is providing optimized estimates of the clear–sky DLI to use then in future works to estimate the cloud direct radiative effect in the infrared spectrum at THAAO, is clearly identified and of potential interest also for other researcher. The work fits within the scope of AMT and publication can be suitable after minor revisions. I propose here some suggestions to potentially clarify some aspects and broad the conclusions, as well as some minor technical remarks.

We thank Dr Di Biagio for the useful comments and suggestions that improve the clarity and effectiveness of the work done and its presentation.

Abstract, line 20: "The bias displays a significant improvement when the coefficients of the different formulas are calculated using the THAAO dataset." This is somehow obvious, I guess. I would suggest: "As expected, the bias displays a significant improvement when the coefficients of the different formulas are calculated using the THAAO dataset."

Done

Abstract, line 21–22: "The presence of two full years of data allows the investigation of the inter–annual variability,". I am not sure that two years of data are sufficient for performing an analysis of the inter–annual variability. On the contrary, as discussed in Sect. 4.1, the two years are interesting to be used in this study since they show different conditions to be tested. I would put this aspect in evidence in the abstract.

The sentence "The presence of two full years of data allows the investigation of the inter-annual variability, and the use of different years for the determination of the coefficients and the evaluation of results" has been changed into "*The presence of two full years of data allows the determination and the applicability of the coefficients for singular years and the evaluation of results*".

Introduction, lines 28–29: references would be good at the end of this first sentence

The following reference has been added to the sentence.

Taylor, P. C., Boeke, R. C. , Boisvert, L. N. , Feldl, N. , Henry, M. , Huang, Y., Langen, P. L., Liu, W., Pithan, F., Sejas, S. A. and Tan, I.: Process Drivers, Inter-Model Spread, and the Path Forward: A Review of Amplified Arctic Warming, Front. Earth Sci., 9, doi: 10.3389/feart.2021.758361, 2022.

Introduction, line 40 and wherever this applies: very minor comment, but when possible it is better avoiding to start sentences with acronyms

Done

Introduction, lines 40–42: references would be good at the end of this sentence

The following reference has been added to the sentence.

Shupe, M. D. and Intrieri, J. M.: Cloud Radiative Forcing of the Arctic Surface: The Influence of Cloud Properties, Surface Albedo, and Solar Zenith Angle, J. Climate, 17, 616–628, https://doi.org/10.1175/1520-0442(2004)017<0616:CRFOTA>2.0.CO;2, 2004.

Section 3: Two comments here:

1/ it would be useful to show (even in the supplementary) a plot illustrating the distribution of clear–sky periods during the two years under consideration. Somehow the discussion of the results in terms of seasonality is missing and it would be good to start from the identification of the distribution of clear sky periods.

We agree with the reviewer that a more detailed description of the annual data distribution could add useful information to the work, but we believe that including this other information in the paper could make the work too long. Moreover, from our perspective, it is not completely within the scope of the article.
Therefore we have prepared additional material presenting both the annual variability of the data and an evaluation of the performance of the parameterizations in winter and summer by applying the formulas in the months of January-February and July-August of 2017 and 2018.

In our opinion, a discussion on the performances of the different parameterizations as a function of seasonality does not fall within the main objectives of the work. Indeed, the purpose of the study is to find parameterizations able to reproduce the DLI starting from the variability of the parameters that most influence it, i.e. T_s, e_s and IWV.
The main goal of this work is to evaluate the effectiveness of these parameterizations throughout the year in the Arctic sites similar to THAAO, i.e. for all those stations that present a distribution of the values of these parameters similar to that reported in Figure 3 of the paper, analyzing which parameterizations best reproduce the DLI values depending on the atmospheric variable used, providing the most suitable coefficients for the purpose and indicating an uncertainty to be associated with the parameterizations used regardless of the year or period of study. Also the comparison between the results found in the two years have this meaning.

We agree with the reviewer that using "only" two years the term "interannual variability" can be misunderstood. On the other hand, we believe that one of the most interesting points of the work is using two different annual databases, obtaining the optimized coefficients for one year and applying them to the other year, verifying that the uncertainty associated with the parameterization (i.e. RMSE value) does not present substantial variations, even if the two years show differences in the evolutions of the atmospheric variables.
In this way, the applicability of the results obtained if applied to different years as regards for THAAO or for other sites that present a meteorological variability comparable to that reported in Figure 3 was verified (at least on two full annual cycle).
The functions used to optimize the parameterizations (see the specific answers to the reviewer comments in the following) tend to minimize the bias between the measured and the parametrized DLI values, obviously depending on the dataset used. For example, in the hypothesis of a field campaign of limited duration, better results would probably have been obtained by optimizing the parameterizations over that time interval (or rather, over an slightly longer interval). Evaluating the monthly or seasonal performance using an optimized annual parameterization has the meaning of evaluating how much the uncertainties associated with a given

parameterization remain constant during the year (for example, in terms of RMSE), without however changing the meaning of the annual results.

Following a different approach to this work, it would have been possible to group the dataset used at a seasonal level and then carry out an analysis similar to the one carried out. By dividing the data in sub-annual periods, like seasons, we would have found performances that were probably slightly better than the annual ones.

On the other hand, the results thus obtained would have presented some limitations including:
1) the difficulty of having several uncertainties to apply during the year.
2) the need to manage the overlapping periods of the different seasons in terms of choosing the most suitable set of coefficients for each parameterization and different meteorological condition.
3) a reduced applicability of the results to other sites.

At the end of the answer to reviewer seasonal variability of cloud free occurrence is presented, together with the performance of the annual parameterizations for the months of January-February and July-August 2017 and 2018, respectively based on the parameterizations optimized on the annual data of 2017 and 2018, which will be also provided as a single attachment (supplementary material).

2/ I understand this cloud–screening procedure is somewhat original. If this is correct, would not be worth to put this more in evidence in the paper? This could imply a few changes in the title/abstract/general discussion. I let the authors to evaluate if this is pertinent or not.

We acknowledge the reviewer for the suggestion. Given the interest also shown by the other reviewer, we are evaluating the hypothesis of dedicating a specific work to the characterization of this cloud screening method which, however, goes beyond the scope of the present work. In fact, to better refine this method it would be necessary to at least add data from a ceilometer and a sky imager in order to have (at least for daytime measurements) independent measurements of the presence of clouds and cloud cover to validate the cloud screening method.

Section 6, general comment: the analysis of the seasonal dependence of the performances of the different formulas is not provided. Have the authors' analysed possible seasonal changes in the performances of the different formulas both for original and the THAAO–optimized formulations? Can they provide some insight on this aspect or derive useful information from this analysis? In some points the discussion the authors mention that original formulations derived based on partial year data are not good because of this limited time period used as reference. In order to better understand this point, it would be appropriate to discuss the seasonal coverage of clear–sky data used for THAAO analysis (see previous comment) and the seasonal changes (if there are) of the performances of the different formulas.

See the answer to the previous comment and the new supplementary material.

Section 6.2: can more details be provided on the procedure to estimate the optimized coefficients?

To estimate the coefficients of the different parameterizations starting from the data measured at the THAAO Observatory, the routines of the IDL software version 7.1.2 were used. As shown in Table 2 of the manuscript, all DLI parameterizations are based on the emissivity parameterization except for ID#13, ID#16, and ID#17; therefore, with the exception of these cases, the emissivity was calculated, and the performances were subsequently obtained by comparing the measured and parameterized values of the DLI.
Most formulations can be linearized as a function of the parameters e_s, T_s, IWV and/or a combination of them, e.g. e_s/T_s; in this case the LADFIT function, which uses a minimum absolute deviation method was applied. In other cases, such as for ID#11, the emissivity was calculated using the POLY_FIT function which applies a polynomial least squares fit. For the parameterization of ID#16 and ID#17 the REGRESS function to

perform a multiple linear regression fit has been used. For the parameterization of ID#7 and ID#12 we applied the CURVEFIT function which uses a gradient expansion algorithm to compute a nonlinear least squares fit. Given the size of the database, the fits were applied without considering any uncertainty in the measurements. Specific details of the used functions can be found on the web page https://www.nv5geospatialsoftware.com/docs/routines-1.html.

Section 7, lines 400–401, I would skip those two lines, it is a repetition

The sentence has been modified to cite the information regarding the number of parameterizations used in the article in the conclusions (now line 421).
It now reads: *"Seventeen different formulas have been chosen to be tested against the THAAO observations"*

Section 7, line 408: see my previous comment on the "inter–annual variability"

The sentence has been modified accordingly.

Section 7, line 432: worth to point out that the 5–7 Wm^{-2} is within the uncertainty of the DLI by pyrgeometer, as discussed earlier in the paper

The sentence at line 456 has been changed into "*Thus, specialized formulas allow to retrieve the clear sky DLI within about 5-7 W/m^2, as also suggested by the distribution of the percentiles, that is of the same magnitude of the uncertainty of the DLI measurements carried out by high quality pyrgeometers.*"

Section 7, general comments: based on the analysis and conclusions of the paper, two points can be raised and maybe deserve a word of conclusion from the authors. Firstly, as at the end the DLI parameterizations are site–dependent, one may ask if, for a specific site, the tuned–parameters derived for a specific year can be used to predict the DLI no matter which year of application or not. In other words, a discussion about which statistics/time range can be reasonable to consider for proper use of tuned–parameterizations should be added. Secondly, does the authors have any hypothesis on why the DLI parameterizations are site–dependent or which parameters seems to affect the most the formulation?

As briefly mentioned in the introduction, line 36-40, the DLI is determined "by the atmospheric concentration of the main greenhouse gases and their radiant temperature". The best estimate of the DLI should be obtained by knowing the vertical distribution of the main greenhouse gases, mainly water vapor, and of the air temperature and using an updated radiative transfer model.
DLI parameterizations are as accurate as they are able to represent the atmospheric profiles, mainly in the lowest layers. Their adaptability to different sites depends mainly on the fact that each site presents different meteorological conditions, not only in terms of variability at screen level, but also in the vertical profile (e.g. temperature inversions).
It is therefore notable that parameterizations developed for global use such as those of Prata, 1996, and Dilley and O'Brien, 1998, provide the best results even for THAAO environmental conditions in both their original and site-optimized versions.

Any conclusion or direction to provide generalization of the results would be for sure very much appreciated by the community. In particular, to what extent the tuned parameters for a site (THAAO for example) can be of relevance for other sites or not can be discussed by the authors.

The following sentence has been added to the conclusions, line 431:

The coefficients of the parametrization optimized for the 2017 and 2018 meteorological conditions at the THAAO (see paragraph 4.1 Dataset characteristics) are provided in Table 2; their performance and therefore their applicability to other sites is mostly related to the meteorological condition. The more they are similar

to THAAO, the more the parameterizations found for THAAO can be directly applied with similar uncertainties (e.g. RMSE value).

SEASONAL VARIABILITY OF CLOUD-FREE OCCURRENCE AND PERFORMANCE OF THE ANNUAL PARAMETERIZATIONS FOR THE MONTHS OF JANUARY-FEBRUARY AND JULY-AUGUST OF BOTH 2017 AND 2018.

Figure 1b shows the monthly availability of the data necessary to carry out the study calculated using the number of minutes contained in each month. Even in the months with more data, the coverage does not exceed 85% mainly due to the measurements of the microwave radiometer that has scheduled interruptions of the IWV measurements to carry out internal calibrations and measurements at different zenith angles necessary to estimate the temperature profile.

As can also be deduced from the time series shown in Figure 1 of the manuscript, the periods with the greatest interruptions concern November and December 2017, caused by the missing of data first of the IWV and then of DLI and IBT, and of September 2018 due to the lack of DLI and IBT data, which are acquired by the same datalogger. In January and December 2017 there is a reduced percentage of observations also due to the snowfall accumulation on the pyrometer or pyrgeometer or both (see paper lines 184-189), which was manually assessed by looking at the data.

The occurrence of cases of clear skies was evaluated both with respect to the total number of minutes present in each month and with respect to the performed measurements; they are called *monthly clear sky* and *normalized monthly clear sky*, respectively.

Using the *normalized monthly clear sky* (discussed in the present work), the strongest annual variations in the clear sky occurrence in the two years are found in April and June-July, months that in both 2017 and 2018 have presented a high number of observations.

The difference observed between April 2017 and 2018 constitutes an anomaly compared to the clear sky occurrence observed in March and June, which are substantially the same in the two years.

The situation is different for the differences shown between June-July 2017 and 2018. In fact, the lower value of the clear sky in June 2017 compared to that of 2018, is compensated by the higher value in July 2017 compared to that of 2018, suggesting a different temporal development of the summer season.

For the remaining months, the *normalized monthly clear sky* is very similar for the two considered years.
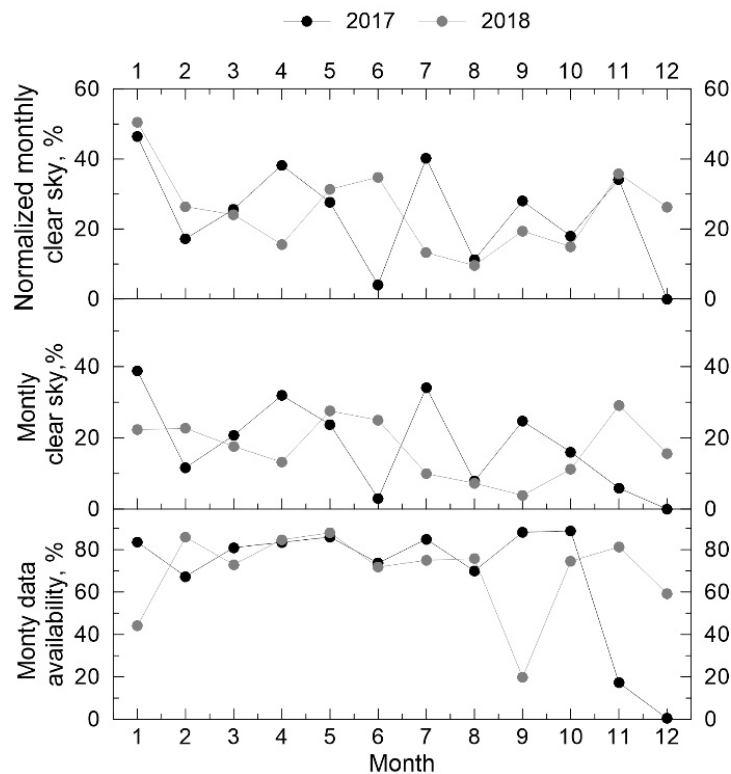

Figure 1b. From bottom to top: time series of monthly data availability, of the monthly occurrence of clear skies calculated considering all the measurable data and of the normalized monthly occurrence of clear skies calculated with respect to the data measured.

Overall, both the monthly availability of the data used and the actual occurrence of clear skies during the two years are homogenous in relation to the variability of the parameters in Figure 3 of the manuscript, showing that the results are affected by reduced monthly representativeness of the measures.

As suggested by the reviewer, the effectiveness of the parameterizations optimized in 2017 and 2018 was tested for different periods of the year. Based on data temporal representativeness, we have chosen to verify the performances by grouping the months of January-February and July-August representative of the winter and summer periods, respectively. To present the results of the annual parameterization applied to different periods of the year, we show the bias and the RMSE values separately, reporting in Figure 2b the performances evaluated on the whole year, on the winter and summer periods, for both 2017 and 2018.

Some particularly high RMSE values (corresponding to parameterizations of ID#3, Swinbank (1963), and ID#6, Brutsaert (1975), in 2017 and of ID#6 in 2018), or bias values very far from 0 (ID#1, i.e. Maykut and Church (1973), ID#3 and ID#6 in both 2017 and 2018), do not appear in Figure 2b because they are very high and representative of parameterizations with low performance as already highlighted in the text of the manuscript. This choice was made to highlight the better performance of the other parameterizations.
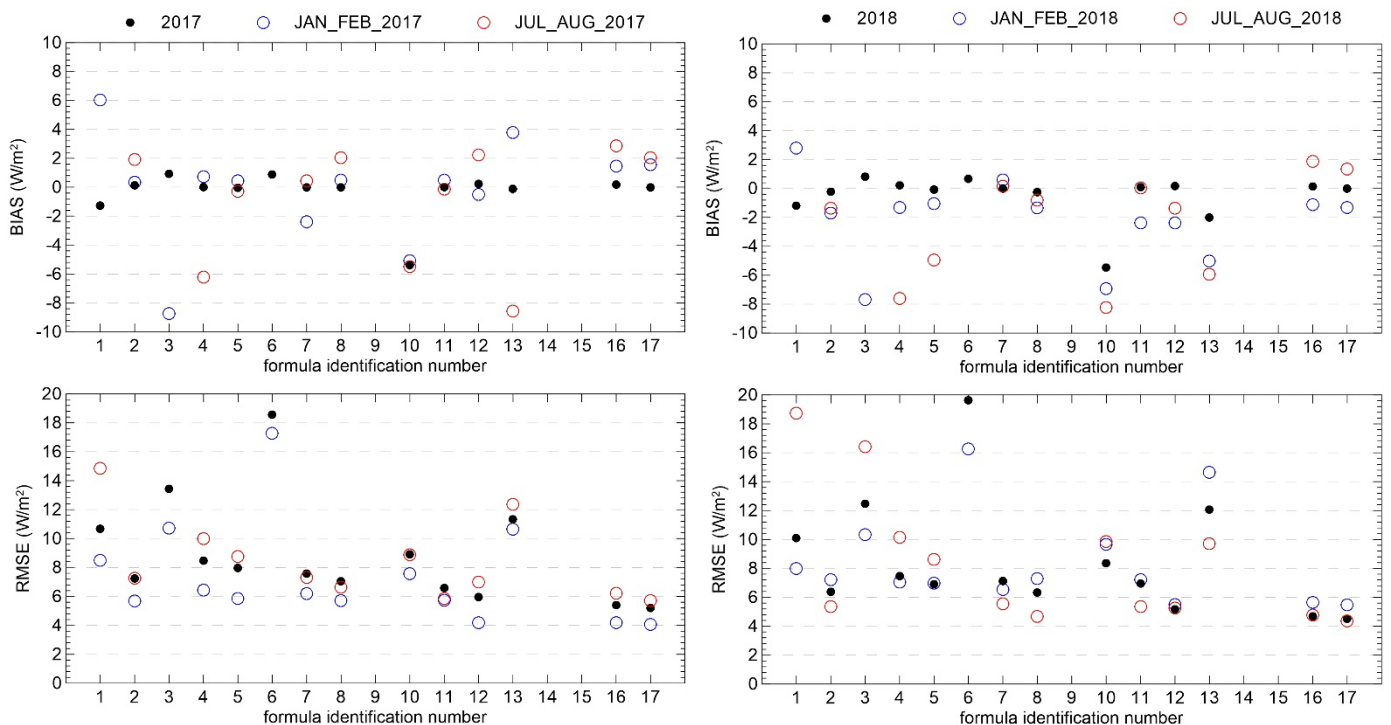

Figure 2b. From bottom to top: on the left the values of RMSE and bias for the annual (black dot), winter (blue circle) and summer (red circle) data of 2017; on the right, the same for 2018. The performances were evaluated against the optimized parameterizations on annual data from 2017 and 2018 respectively. Some values of the RMSE of the parameterizations ID#3 (Swinbank, 1973), ID#6 (Brutsaert, 1975) in 2017 and of ID#6 in 2018, as well as some values of the bias of ID#1 (Maykut and Church, 1973), ID#3 and ID#6 in both 2017 and 2018, are out of scale of the plot (see comment in the text).

In terms of bias, considering both the seasonal differences of 2017 and 2018, no single parameterization always presents the best performance. In 2017 both ID#5, Ohmura (1981), and ID#11, Jin et al. (2006), show very small summer and winter values close to the annual ones. On the other hand, in 2018, ID#7, Satterlund (1979), shows this behavior, while ID#11 shows practically zero summer and annual bias, but a winter bias slightly larger than 2 W/m^2. The parameterizations of ID#12, Prata (1996), ID#13, Dilley and O'Brien (1998), and ID#14, Dilley and O'Brien (1998), have annual bias very close to 0 and present respectively positive and negative biases in the summer and winter seasons (ID#12 in 2017, and ID #16 and ID#17 in 2018), or only negative for ID#12 in 2018 or only positive for ID#16 and ID#17 in 2018.

Regarding the seasonal bias, there is no homogeneous behavior between the parameterizations.

Considering the RMSE, in 2017, there is a behavior shared by practically all stations: the winter RMSE values are lower than the annual ones, and the summer ones are approximately equal to or slightly higher than the annual ones. It is interesting to note that in 2018 the RMSE values do not present this feature, for various parameterizations (ID#2, ID#7, ID#8, i.e. Idso (1981) and Andreas and Ackley (1994), ID#11, ID#12, ID#16 and ID#17) the summer value being less than or equal to the annual one.

The parameterizations ID#11, specifically developed for the Arctic and using both the values of e_s and T_s, ID#12, ID#16 and ID#17, generally show the lowest values of RMSE. Although the ID#12, ID#16 and ID#17 parameterizations use the values of e_s, T_s and IWV, it is notable that having been formulated for a global application, obtain these good results when applied to the peculiar conditions of the arctic region.

However also parametrizations ID#2, ID#7, and ID#8, show very good performance using respectively only e_s, ID#2, or e_s and T_s (ID#7 and ID#8).

The best performances are maintained around RMSE values between 4 and 6 W/m^2 and, therefore, are statistically comparable with the accuracy of the pyrgeometer measurement.

In conclusion, we want to highlight that the RMSE is calculated as the quadratic sum of the bias and the standard deviation (see paragraph 4.2 of the manuscript), therefore parameterizations that present a non-negligible bias (e.g. ID#16 and ID #17) and low RMSE have a lower standard deviation, indicating the ability to better follow the variability of the measurements, compared to the parameterizations showing a lower bias, but higher RMSE values.