Dear editor,

Thank you for your efforts to help with the editorial process with our manuscript. We have organized our response as follows. First, we include the responses to the referees. Then, we include a track changes version of the manuscript with all removed text in red and new text in blue. The goal was to answer each reviewer comment and integrate the related changes into the revised manuscript (attached).

Thanks in advance for your continued work as editor of this manuscript and I look forward to hearing from you in the coming weeks.

Reviewer #2:

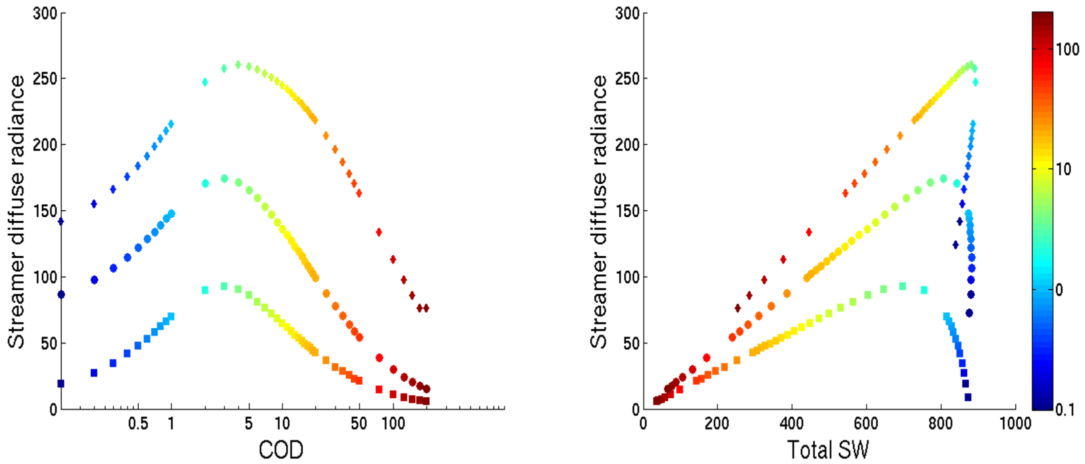
AC : The authors would like to thank the Reviewer#2 for his/her careful review of our manuscript. We addressed each comment individually and have revised the manuscript accordingly.

Overview - This paper documents an initial evaluation of whether the background noise from the lidars installed on IAOOS drifting buoys in the Arctic Ocean can be used to estimate at least some of the components of the surface radiation budget. The approach relies upon combination of the lidar noise measurement combined with information from a radiative transfer model, along with some input from ERA5 reanalysis fields. The method progresses through a number of steps, some of which depend upon assumptions about conditions. The impact of these assumptions is only explored for some of them, and the uncertainties introduced by the others remain unquantified. While well written overall, the method described is complicated, and there are a few places where the discussion becomes confusing and would benefit from being clarified – this often results from a term being used before it is defined – these are noted below. There are also a number of grammatical slip-ups and typos, detailed below.

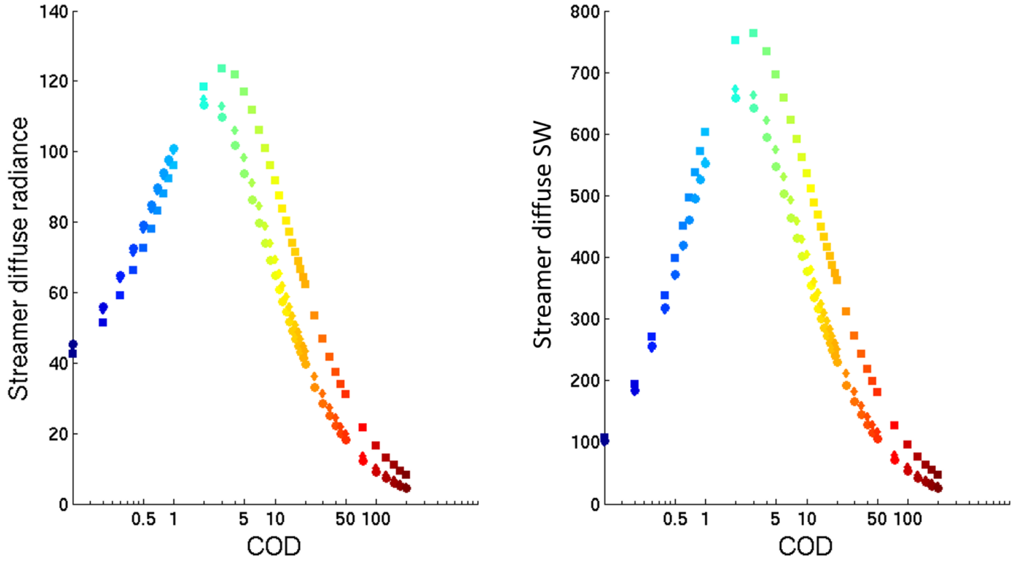
Major comments - While the approach documented appears to work remarkably well, I have some concerns about the potential for self-correlation to give a false sense of its effectiveness a result of the use of the same data points both for calibration and evaluation of the lidar data. The entire approach rests on the determination of the relationship between the lidar noise, B, and downwelling scattered shortwave radiance for the lidar's (narrow) field of view, Ls↓. This is instrument specific and determined by fitting a function to measurements of B and Ls↓, ideally – as noted in the paper – this would be done for clear sky conditions. The authors note that there are very limited periods with clear skies in the data set, and they opt instead to use a slightly less direct approach. They use a value of Ls↓ estimated for cloudy conditions, with Ls↓ derived from a radiative transfer model, STREAMER, with a cloud optical depth estimated from the directly measured total downwelling shortwave irradiance, FSW↓. This is shown in figure 4. On the basis of this figure alone, which shows a very strong linear relationship, I think the proposed technique is almost certainly useful. However, when the full technique is implemented: lidar noise is used to estimate Ls↓, which is used with the STREAMER results to estimate cloud optical depth and FSW↓, the final results are evaluated by comparison with the same N-ICE data used to calibrate the lidar in the first place, and with the same assumptions in place about other quantities that might affect the results:

surface albedo, cloud thickness, and cloud microphysical properties. While it is clear that there is a strong relationship between B and Ls↓, if some bias is introduced into the calibration because of an invalid assumption – use of a fixed albedo (perhaps too high or low), cloud microphysical properties, cloud thickness – which might affect either the gradient, KL, or offset, b, of the linear fit, then that bias CANNOT be identified by the evaluation used here. The error analysis provided in the appendices deals effectively with random errors – the uncertainty deriving from inherent uncertainty in the measurements or assumed values, but not with any potential mean bias that might arise from the initial calibration.

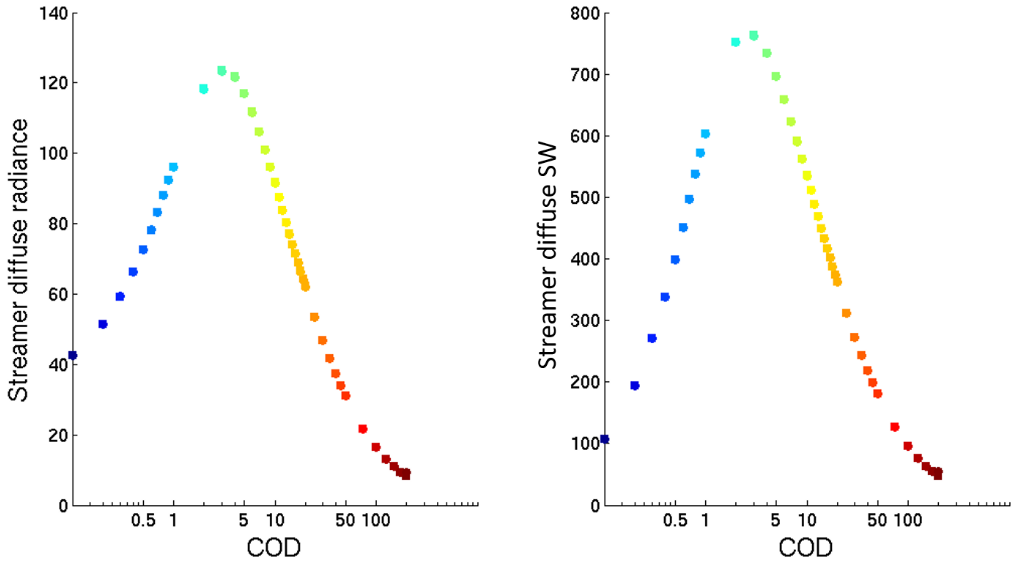
We thank the reviewer for this detailed useful comment underlying the limits of our approach. We agree with the reviewer, some bias can be introduced into the calibration because of an invalid assumption into STREAMER. We did actually take into account those possible biases and considered assumptions to limit these biases. We calculated the uncertainty on the optical depth, FSW and Ls↓ (shown in appendix) for different assumptions on the values of the albedo, cloud microphysical properties, cloud thickness and cloud height in STREAMER. New Figures are reported below.



Ls↓ as a function of COD (left) and FSW (right). The cloud base is set-up at an altitude of 1 km and its thickness is 1 km. The modelled cloud is entirely composed of liquid water droplets of radius 5 µm and εs is 0.95. The different curves correspond to different albedo (α) values: 0.9 (diamonds), 0.5 (circles) and 0.05 (squares). The color scale follows the evolution of the optical thickness.



Ls↓ (left) and FSWs↓ (right) as a function of optical thickness. The cloud base is set-up at an altitude of 1 km and its thickness is 1 km. The surface parameters are the following: ε_s is 0.95 and α is 0.9. The different curves correspond to different hydrometeors :cloud composed of water droplets of radius 5 µm (squares), cloud composed of ice crystals as solid columns of radius 5 µm (circles), cloud composed of ice crystals as disks of radius 5 µm (diamonds). The color scale follows the evolution of the optical thickness.



Ls↓ (left) and FSWs↓ (right) as a function of optical thickness. The cloud has a thickness of 1 km. The cloud is entirely composed of liquid water droplets of radius 5 µm. The surface parameters are the following: ε_s is 0.95 and α is 0.9. The different curves correspond to different cloud base altitudes; 1 km (squares), 5 km (circles), 10 km (diamonds). The color scale follows the evolution of the optical thickness.

A wrong assumption of the albedo would be responsible of large discrepancies on the results of the approach. We have decided to use measured albedo during N-ICE to

constraint STREAMER rather than a fixed one, assuming that the albedo values at the locations of the buoy and the N-ICE ice camp are similar. The induced uncertainty is discussed in Section 4.4 . The assumption on the cloud height and thickness have a negligible influence because the optical depth is fixed. The assumption of the cloud microphysical properties may lead to a small bias and it was already taken into account in the uncertainty of the optical depth and FSW.

We added in the new manuscript the uncertainty on the Kl constant caused by the assumption made in STREAMER, and the new sentences are

"The uncertainty in KL due to the choice of cloud properties in STREAMER is not significant (ΔKL = 0.38 W−1 m2 sr). But the influence of a change in the surface albedo $\Delta\alpha$ is significant (ΔKL = 5.4 W−1 m2 sr). To reduce uncertainties, the albedo obtained from the N-ICE measurements is used to constrain STREAMER runs. The spectral dependency of the solar scattered radiance calculated in the range 200 −3600 nm is implicitly embedded into this slope KL. "

I appreciate the problem the authors face – they are trying, after the fact, to derive quantities that were never intended to be measured by this instrumentation, and lacking some of the support measurements that one would make if planning this prior to the field campaign. They're done a good job, but could perhaps address the problem of potential calibration bias more effectively. The available clear sky data is very limited, but it is not zero – May 23 is stated to have 24 hours of clear skies. Do estimates of B and $Ls\downarrow$ from clear sky conditions – however limited – fit the function derived from the cloudy cases?

Clear sky conditions are indeed present on May 23, but unfortunately there is no lidar observation at this time. As a consequence, it is not possible to estimate the relation between B and $L_{s\downarrow}$ from clear sky conditions for the studied period.

The period used here, is a 44 days, but only 18 discrete lidar measurements are used, each a 10 minute average, from ~4 measurements per day, so about 10% of the 176 total measurements. While the method clearly has merit, does this very sparse data set imply it's operational use would be likely to return similarly sparse data?

This is one limitation of our study; in spring, there are not 176 measurements available, but only 65. There are up to 4 measurements a day. Out of those 65 discrete values, only 18 are actually used (28% of the total measurements) because of the icing issues detailed in the paper. The rest has been discarded from the analysis to prevent any potential bias. The heating system onboard the IAOOS buoys has been improved after this campaign; hence a smaller number of lidar observations are biased by icing issues, enabling a more extensive dataset to be used over the six years. In this study however, we are focusing on this sparse dataset, as the comparison against co-localized NICE observations is crucial to assess the performance of the approach.

Line 30: the authors state "clouds cover up to 80% of the region at all time and are primarily composed of low-level mixed phase clouds". This true for the summer but not necessarily for the winter season when low level clouds are less frequent.

We agree with the reviewer that low level clouds are less frequent in winter. We replaced 'all time' by 'spring and summer' in the new manuscript.

Line 64: regarding the reference to MOSAiC. Keep an eye on https://online.ucpress.edu/elementa/searchresults?fl_SiteID=1000091&page=1&tax=231 –the initial programme overview papers are currently in press, and should be available by the time this paper is accepted.

We thank the reviewer for this updated list of in press papers. The atmospheric reference have been added in the revised manuscript.

"Shupe, M. D., Rex, M., Blomquist, B., Persson, P. O. G., Schmale, J., Uttal, T., Althausen, D., Angot, H., Archer, S., Bariteau, L., Beck, I., Bilberry, J., Bucci, S., Buck, C., Boyer, M., Brasseur, Z., Brooks, I. M., Calmer, R., Cassano, J., Castro, V., Chu, D., Costa, D., Cox, C. J., Creamean, J., Crewell, S., Dahlke, S., Damm, E., de Boer, G., Deckelmann, H., Dethloff, K., Dütsch, M., Ebell, K., Ehrlich, A., Ellis, J., Engelmann, R., Fong, A. A., Frey, M. M., Gallagher, M. R., Ganzeveld, L., Gradinger, R., Graeser, J., Greenamyer, V., Griesche, H., Griffiths, S., Hamilton, J., Heinemann, G., Helmig, D., Herber, A., Heuzé, C., Hofer, J., Houchens, T., Howard, D., Inoue, J., Jacobi, H.- W., Jaiser, R., Jokinen, T., Jourdan, O., Jozef, G., King, W., Kirchgaessner, A., Klingebiel, M., Krassovski, M., Krumpen, T., Lampert, A., Landing, W., Laurila, T., Lawrence, D., Lonardi, M., Loose, B., Lüpkes, C., Maahn, M., Macke, A., Maslowski, W., Marsay, C., Maturilli, M., Mech, M., Morris, S., Moser, M., Nicolaus, M., Ortega, P., Osborn, J., Pätzold, F., Perovich, D. K., Petäjä, T., Pilz, C., Pirazzini, R., Posman, K., Powers, H., Pratt, K. A., Preußer, A., Quéléver, L., Radenz, M., Rabe, B., Rinke, A., Sachs, T., Schulz, A., Siebert, H., Silva, T., Solomon, A., Sommerfeld, A., Spreen, G., Stephens, M., Stohl, A., Svensson, G., Uin, J., Viegas, J., Voigt, C., von der Gathen, P., Wehner, B., Welker, J. M., Wendisch, M., Werner, M., Xie, Z., and Yue, F.: Overview of the MOSAiC expedition: Atmosphere, Elementa: Science of the Anthropocene, 10, https://doi.org/10.1525/elementa.2021.00060, 2022"

Line 54: "The surface cloud radiative effect is therefore positive from September to April-May and negative in summer" – it should perhaps be acknowledged that this is also a function of latitude.

Our intention was to mention the Svalbard region, where the campaign took place. We agree with the reviewer that this is not the case everywhere in the Arctic. We have modified the text as follows : " The surface cloud radiative effect in the Svalbard region is positive from September to April-May and negative in summer (Walden et al., 2017; Ebell et al., 2020). Near Greenland, the cloud radiative effect is positive all year round (Miller et al., 2015). "

Line 159-161: The cloud used for the STREAMER simulations is defined here, and has a fixed altitude and depth. I appreciate that for a given optical depth this probably doesn't greatly affect the results, but I wondered why these weren't height and depth were not taken directly from radiosonde profile – at least for the calibration and to evaluate the range of any impact. Using a better estimate of the actual cloud properties for the calibration of lidar noise against scattered radiance, rather the fixed values, which are then also used for the determination of optical depth from lidar noise, would eliminate one of the closed-loops when validating the method against the same N-ICE data used for the calibration.

We tested the uncertainty caused by a wrong assumption of the altitude and depth of the cloud in STREAMER on the optical depth, Ls, Fsw and Kl. Little to no change is observed by changing the altitude and depth of the cloud in STREAMER because the

optical depth is given. Instead of using a fixed value of altitude and depth, we are now using the average altitude (400 m) and depth (800 m) determined from the IAOOS lidar measurements. However, it did not lead to a significant change on the results. We have modified the text as :

"Based on the IAOOS lidar's measurements, clouds have a mean geometrical depth with base and top at 400 and 1200 m above mean sea level, respectively. These clouds are assumed to be composed of two cloud layers: a 150 m-width cloud layer composed of 6.9 ±1.8 µm-diameter water droplets overlying a 650 m-width cloud layer composed of 25.2 ±3.9 µm-diameter hexagonal ice crystals, as described by McFarquhar et al. (2007) in similar conditions."

Figure 4: It would be useful if on the right hand panel, a line were overplotted showing the location of the peak in Ls as a function of θ - it would help make clear the ambiguity in τ.

The location of the peak in Ls as a fonction the SZA has been added accordingly in Fig. 4.

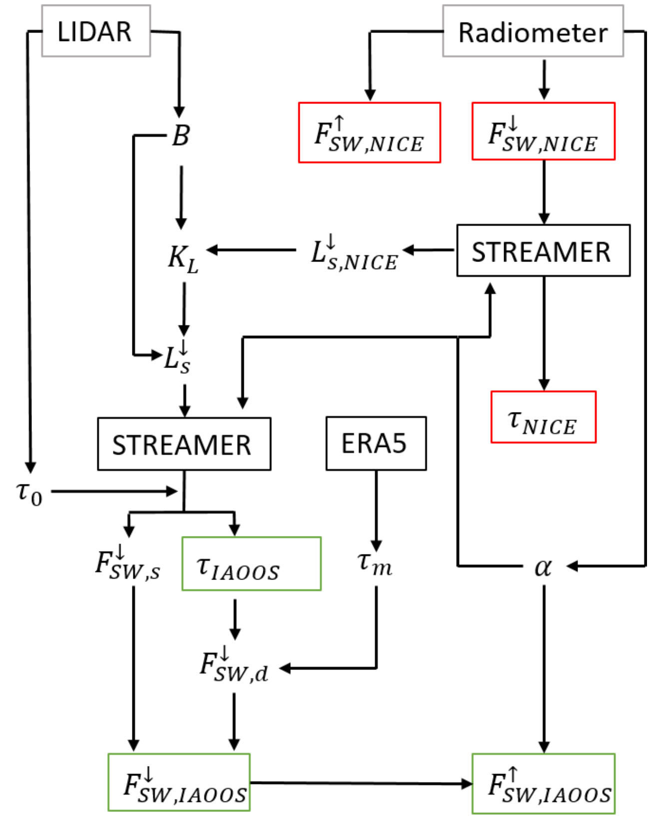
Figure 5: the caption states that the results here are plotted for 6 values of τ (the different coloured points) and 'various values of θ' – I assume that at a given τ each point corresponds to a different value of θ, but that is not obvious from the figure or caption.

This is true. We have mentioned that each point corresponds to a different value of θ in the caption of figure 5.

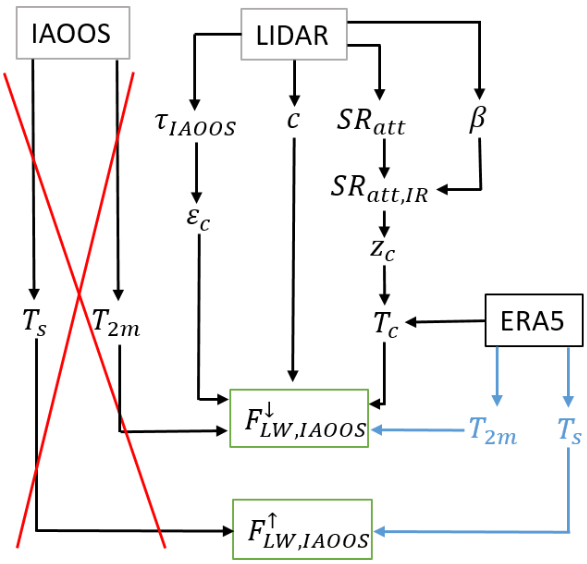
Lines 190-194: this brief description is the closest thing given to a complete description of the method proposed. The various components are covered in more detail in various sections, but it would be useful to give a clear, step by step, breakdown of the full method. Here the description is rather vague, e.g. "...B derived from the lidar helps to determine..."

According to the comments posted by Reviewer#1, we have completely reorganized the paper. The main cause of the confusion in the submitted manuscript was the fact that the approach was spread over two sections intermixed with some of the results. In the new manuscript, we choose a more classical approach : Section 3 presents the methodology and Section 4 the results. As a consequence, the methodology to get the cloud optical depth from, on the one hand, SW flux measurements during the N-ICE expedition and from, on the other hand, the background signal B of the IAOOS lidar, are all included in Sect. 3. This latter also describes the method to obtain the regression slope Kl required to convert B to a scattered radiance. Every step of the approach is now presented in Section 3.2 for the SW flux estimation and Sect. 3.3 for the LW flux estimation. All the results of the determination of those variables have been moved to Section 4.1 (previously Section 4.4).

We have also added two schematics (reported below) describing flowcharts to clarify the article structure, the notations used in the paper and to help the readers follow the different steps, which observations are exactly used and where.



Flowchart describing the method to get the downward SW flux $F\downarrow_{SW}$ from the solar background B measured by the lidar. The estimated optical depth τ_{IAOOS}, $F\downarrow_{SW}$ and $F\uparrow_{SW}$ in green are compared to their observed counterparts from N-ICE measurements in red. τ_{IAOOS} is then used to estimate the downward LW flux.



Flowchart describing the method to gt the downward LW flux $F\downarrow_{LW}$ from the observations onboard the IAOOS buoy (temperature and lidar measurements). Ts and T2m are over this period (April-June 2015) are instead taken from ERA5 (blue) because there are not available from the IAOOS buoy due to instrumental issues (red cross). τ_{IAOOS} is obtained from the flowchart describing the retrieval of SW fluxes.

Line 195: "is simply obtained from the equation detailed by Minnis et al. (1993)." – Minnis et al.(1993) has a lot of equations, none of which perfectly match this one. I think you refer to equation 21, but please cite the intended equation explicitly.

We indeed refer to equation 21 from Minnis et al. (1993) and equation 29 from Minnis et al. (1993). They have now been cited explicitly in the text.

Equation 5: the term εc is undefined...until 10 lines below. And the term c – the cloud mask, isn't defined for 5 lines. It would help the reader if all terms were defined immediately after the equation, rather than much later in the discussion.

We define now the terms εc and c earlier, just after equation7.

I'm not sure I understand the full reasoning for equation 5 and its relationship with Minnis eqn 21. Two questions: 1) In the cloudy case, eqn 5 gives the downwelling LW irradiance as the sum of those for the cloud (emitted, from cloud at temperature Tc) and the surface (reflected, surface temperature Ts). Minnis eqn 21 looks like the same equation: FLW = ∈B(Tc) + (1-∈)B(Ts) , where B(T) = σT4 but it addresses a slightly different situation, the upwelling radiation seen from a satellite, with cloud temperature Tc and Ts the 'clear scene' (surface) temperature. Here the radiation from the surface is seen through the cloud rather than being reflected. Why the difference for your case? 2) in the clear case you give the downwelling LW radiation simply as that emitted by the lowest level of the atmosphere (at 2m). I don't understand how that is reasonable, why no contribution from higher levels (or the background of space?) – the atmosphere is largely transparent at the wavelengths concerned here or the measurement of cloud temperature wouldn't be possible.

1) We apologize for this. There was a mistake in our equation 5 (6 is the new manuscript): the upward LW radiation emitted from the surface is either absorbed by the could layer (fraction εc) or transmitted to space (fraction 1- εc). The reflected fraction from the cloud can be neglected in the infrared spectrum. We have corrected the equation. Fortunately this had close to no effect on the results since the optical depth of the clouds were strong (εc ≈ 1)

2) As mentioned by the reviewer, the downwelling LW radiation is written as that emitted by the lowest level of the atmosphere (at 2 m), as if the atmosphere was a grey body emitting at 2 m under clear skies only, with a broadband IR emissivity of 0.7. This proxy was used by Shakespeare and Roderick (2021), Zhang and Zhang (2001) and Brutsaert (1975) under the assumption of a nearly standard atmosphere.

"Brutsaert, W.: On a derivable formula for long-wave radiation from clear skies, Water Resources Research, 11, 742–744, https://doi.org/10.1029/wr011i005p00742, 197

Shakespeare, C. J. and Roderick, M. L.: The clear-sky downwelling long-wave radiation at the surface in current and future climates, Quarterly Journal of the Royal Meteorological Society, 147, 4251–4268, https://doi.org/10.1002/qj.4176, 202

Zhang, X. and Zhang, J.: Heat and Freshwater Budgets and Pathways in the Arctic Mediterranean in a Coupled Ocean/Sea-ice Model, Journal of Oceanography, 57, 207–234, https://doi.org/10.1023/a:1011147309004, 200"

Line 253: here it is stated that there are 20 points used to calibrate the lidar noise B against the scattered downwelling radiance. I count 18, both on figure 7 and again on figure 6 where the lidar derived cloud optical depth is plotted along with that from N-ICE.

This is a typo; there are indeed 18 points. This has been corrected.

Line 380-381: the authors state that the fact that all LW irradiances were > 230 W m-2 suggests that the lidar detected clouds for the entirety of the spring period. Isn't this forced by the fact that the data used to find the irradiances are selected based on the calculated optical depth?

Our wording was unclear. The relative large values of the LW irradiances (> 230 W/m2) are due to the fact that the IAOOS lidar did not operate under clear skies. The sentence has been rewritten as follows:

"The fact that all N-ICE LW irradiance measurements reported on Fig. 11a are larger than 230 W m−2 confirms that clouds are seen from the N-ICE ice camp at the same instant as the IAOOS buoy lidar measurements."

Line 419: 'The skin temperature taken from the ERA5 reanalysis may be colder than the actual skin temperature' – recent evaluation of the ECMWF IFS model in forecast mode, but essentially the same model used to generate ERA5, has a warm bias of ~1K in the skin temperature, at least during the late summer and early autumn. See: Tjernström, M., G. Svensson, L. Magnusson, I. M. Brooks, J. Prytherch, J. Vüllers, G. Young, 2021: Central Arctic Weather Forecasting: Confronting the ECMWF IFS with observations from the Arctic Ocean 2018 expedition, Quart. J. Roy. Meteorol. Soc. doi:10.1002/qj.3971

We have no fundamental disagreement with the reviewer. In this study, the only possible reason of the discrepancy on the LW fluxes between the ERA5 reanalyses and observations is a cold bias on ERA5. That latter does not necessarily contradict the results from Tjernström et al. (2021). There might be cold bias in late spring/early summer as the sea-ice is melting and a warm bias during the late summer and early autumn associated to the refreezing of sea-ice. The interesting findings of this paper have been quoted in the new manuscript:

"Tjernström et al. (2021) found warm bias of the skin temperature of ERA5 of ~1 K during the late summer and early autumn. Ourstudy does not necessarily contradict those results. There might be cold bias in late spring/early summer as the sea-ice is melting and a warm bias during the late summer and early autumn associated to the refreezing of sea-ice."

Grammar, typos, etc

We thank the reviewer for his careful reading. All grammar and typo errors have been corrected.

Line 10: 'enables to estimate' -> 'enables us to estimate'

Done

Line 21: 'twice as fast then the rest...' -> 'twice as fast as the rest...'

Done

Line 27: 'contribute to regulate the...' -> 'contribute to the regulation of...'

Done

Line 37: "The underdetermined knowledge on the thermodynamical and radiative feedbacks.." – awkward phrasing, better: "The limited knowledge of the thermodynamical and radiative feedbacks..."

Done

Line 39: "the radiative budget, which is the primary source in the surface energy..." -> "the radiative budget, which is the primary contribution to the surface energy..."

Done

Line 71: 'Buoys have also...' -> 'Buoys also have...'

Done

Line 75: 'fluxes from buoys lidar data' -> 'fluxes from the buoys' lidar data'

Done

Line 80: 'to derive both optical depths and radiative irradiances' -> 'to derive both optical depths and irradiances'

Done

Line 82: 'irradiances measurements at the vicinity of' -> 'irradiance measurements in the vicinity of'

Done

Line 91: 'into the pack ice during several months' -> 'into the pack ice for several months'

Done

Line 92: 'tacked' -> 'tracked'

Done

Line 114: 'at a ice camp' -> 'at an ice camp'

Done

Line 121: ‘bandwidths, respectively’ -> ‘bandwidth, respectively’

Done

Line 129: ‘In springtime, Walden et al. (2017); Cohen et al. (2017) indicated’ -> ‘In springtime, Walden et al. (2017) and Cohen et al. (2017) indicated’

Done

Line 134: ‘on 1 minute resolution’ -> ‘at 1 minute resolution’

Done

Line 137: ‘of its emission’ -> ‘of its emission of LW radiation’

Done

Line 139: ‘interpolated at the buoy location’ -> ‘interpolated to the buoy location’

Done

Line 151: ‘same bandwidth as the one of the pyranometer’ -> ‘same bandwidth as that of the pyranometer’

Done

Line 153: ‘by Merkouriadi et al. (2017); Granskog et al. (2018) during’ -> ‘by Merkouriadi et al. (2017) and Granskog et al. (2018) during’

Done

Line 161: ‘water droplets overcoming a 500 m-width cloud layer’ -> ‘water droplets overlying a 500 m-width cloud layer’

Done

Line 169: ‘and decreases afterwards’ -> ‘and decreases above this value’

Done

Line 238: ‘corrected of the Earth-Sun distance’ -> ‘corrected for the Earth-Sun distance’

Done

Line 279: ‘provide a rough information on...’ -> ‘provide a rough idea of’

Done

Line 287: ‘observed the 16 May’ -> *‘observed on the 16 May’*

Done

Line 289: ‘whole set N-ICE...’ -> *‘whole set of N-ICE...’*

Done

Line 474: ‘A1 Uncertainties on τ’ -> *‘A1 Uncertainties in τ’*

Done

Line 496: ‘A2 Uncertainties on...’ -> *‘‘A2 Uncertainties in...’*

Done

Line 509: ‘only depends the uncertainty...’ - *‘only depends on the uncertainty...’*

Done