The manuscript “Radiative fluxes in the High Arctic region derived from ground-based lidar measurements on board drifting buoys” by Loyer et al. describes a method for calculating broadband radiative SW and LW fluxes from IAOOS buoys featuring lidars. The methodology is built around IAOOS buoys that were co-located with comprehensive surface and upper-air meteorology, and radiometric observations collected during the N-ICE2015 campaign. Pairing the buoys with N-ICE2015 and using the campaign to demonstrate the approach within the environment where the approach will be used is a good experimental design and my feeling is that there is a publishable study here. However, this study, or perhaps this manuscript, is not mature enough for publication in its present form. I hedge here because the first issue that needs to be addressed is the organization and clarity of the text (see below). There are also grammatical errors, awkward phrasing, run on and fragmented sentences, and unnecessary subjective qualifiers (e.g., “rather”, “fairly”, “mostly”, “quite”) throughout. It is possible that if the study were better communicated, I would find it more convincing.

A complete rewrite of Sections 2, 3 and much of 4 is needed. There is no logical flow to the narrative as it stands. To begin, explain what steps are needed to calculate fluxes from the buoy, which are currently found spread across all three sections. For example, some necessary introductory information isn’t found until deep in Section 4 (Lines 333-337 & 344-349). Explain what information you have from the buoy and what information you need to get from other sources. Maybe make a figure with a flow chart to help readers follow the methodology that begins with lidar backscatter and ends with a flux. Even after reading the manuscript, I still don’t really understand the purpose of what appears to have been the development of a lookup table using STREAMER in 3.1. I also couldn’t understand why so much effort went into trying to make broadband radiometers produce data that is suitable for deriving from lidar (optical depth) because the necessity of doing this was unclear and because there was a micropulse lidar (MPL) deployed on R/V Lance by the US DoE ARM program, which is better suited for comparison. Optical depths were also sometimes calculated using radiosondes instead of being measured by the IAOOS lidar and soundings were used in some cases while reanalyses were used in other cases. By the end of Section 3, I was not even sure how the buoy data was contributing anything of value to the SW flux calculation because Eqs. (3) and (4) were never
reconciled. The equations for the LW calculation connected better, but if I understand correctly, the only value the lidar is providing was cloud emissivity. The blended use of buoy, ERA5, and N-ICE2015 observations leaves me wondering how you propose to apply this method beyond N-ICE and how the uncertainty will degrade when you don’t have N-ICE observations to incorporate. And ultimately/most importantly, does the lidar add enough useful information to beat ERA5 estimates of flux?

I have a number of additional questions referenced to line numbers below.

Comments:

Can you provide some information on the quality of the lidar observations besides the icing issue? There must be all sorts of challenges with level and signal-to-noise, etc.

Lines 53-55: The summertime values do vary and the cycle is not precisely the same everywhere. An extreme example is Greenland, which is positive year-round (https://www.doi.org/10.1175/JCLI-D-15-0076.1)

Lines 56-73: Something is missing here. I understand there are new buoys in the water, but there is no description of what these buoys are measuring.

Section 2.1. Precisely which data is used in the study is unclear. Fig. 1 presents 3 buoys but the caption says only one is used. This contradicts Line 97, which says data from January-June is used, at least “mostly” (?). The end of the paragraph indicates only April-June data is used, and at that only a subset.

Line 106: “The approach...” is out of place. If this is described later, just remove this sentence.

Section 2.2. Some of the information from 2.1 belongs here and not in 2.1.

Line 22: Note that 5 Wm2 and 3% are equivalent. It’s more like which ever is larger, 5 Wm2 or 3%, and that 3% is 15 Wm2 at 500 Wm2 SWD.

Line 124: What is the purpose of this 70% threshold? Why not use all good data that correspond in time to lidar profiles?
Line 125: How do you define/did you determine what is "too far"?

Figure 2: I understand why you use a negative sign for upwelling data in the top two panels, but it is conventional to use all positive signs, and additionally, as plotted the message using the sign and the arrow nomenclature is somewhat redundant and thus cancelling: i.e., isn’t a negative of the upward arrow a downward arrow? If you make the plot using positive sign conventions it will be much easier to read.

Line 130: By stable do you mean static stability of the atmospheric boundary layer or do you mean that the meteorology did not change much during the period of interest?

Line 131: As a consequence of what?

Line 133: Is a sample of clear skies important for some reason?

Line 134: Would you really classify 550 Wm\(^2\) on a clear day as extreme?

Line 139: Note that this not the native resolution of the ERA5. It is already and interpolated product.

Line 140: Were the soundings from N-ICE2015 sent to GTS and assimilated by ERA5?

Line 150: What is that FOV?

Line 200: I don’t understand, you are using T2m from ERA5 but earlier you said the buoys measure that.

Lines 250-252: I don’t follow. I feel like you have done a clear sky calculation with STREAMER and are using that to estimate K instead of an observation under clear skies. Is that what you mean?

Eq. 11: I don’t follow. This equation contradicts Eq. 4. If the implicit assumption of Eq. 4 is
that b should be 0 then in creating Eq. 11 you have made the assumption that uncertainty in the regression is entirely found in the y-intercept, but it could just as easily be attributed to an error in the slope. Your calculation of K depends on your interpretation of the uncertainty in the regression. If you set b=0 and use the same slope, how much does K change and is that change negligible for the purposes of this study?

Line 289: At this stage, I still haven’t figured out why we are deriving tau from the N-ICE radiometers. What will this accomplish? If this is simply to be able to characterize the limited number of samples from the buoy within the more continuous time series from N-ICE then some re-organization of the paper is needed to provide a more logical progression of the steps. I also wonder why you aren’t using the micropulse lidar from US DoE ARM, which was installed on R/V Lance? That would make a much better comparison data set for the buoy, wouldn’t it, for example avoiding the FOV issues (e.g., Line 315).

Lines 325-331: I’m a little confused about the effect of the frost. Obviously, if it is attenuating then there is no signal. But if there is signal, is the problem that the partial attenuation from thin frost coverage implicitly propagating into B?

Line 398: I don’t understand how upwelling fluxes are coming from IAOOS. The buoy doesn’t measure anything relevant for that that you have described. These values are taken from ERA5, interpolated to the position of the buoy?