Reply to Reviewer #2:

We thank the reviewer for the time and efforts she/he spent reading our manuscript and providing valuable comments. Please find below a discussion of the reviewer's comments (italic). Changes/additions made to the text are underlined and given in quotes.

General comment

I this paper, the Arctic surface albedo simulated with the coupled regional climate model HIRHAM-NAOSIM is evaluated with aircraft and surface-based observations collected during several field campaigns. The study is very relevant for the polar modelling community, the applied method is convincing, and the observational dataset used for the model validation is outstanding. However, I have few major concerns:

1. *The text in Sect 3 and 4 is hardly readable, the expressions are unclear, the language is not suitable for scientific publication and needs to be extensively rewritten. In my detailed comments I only point to few examples, but almost all the sentences require improvement.*

   We have revised the text considerably according to the reviewer's comment, and have rephrased the relevant text passages. We hope that this improves the readability. For details please refer to the detailed comments below.

2. *In some cases, the interpretation of the results needs to be deepened (see my detailed comments). Some results depend on the selected regions and time of the year and cannot be generalized (such as the relative impact of clouds or albedo biases on the bias in surface net irradiance).*

   Please read our responses to the detailed comments below. They refer to that general remark.

3. *In my view, one of the most striking results is the model underestimation of surface albedo after the onset of melting (Fig 7). The underestimation is explained as due to the fact that, when snow disappears, the ice surface is represented as bare ice and not as the surface scattering layer that forms during the melting. This result deserves more discussion.*

   We followed the Reviewer's suggestion and added a discussion of the issue associated with the missing surface scattering layer, and we included the reference to Macfarlane et al. (2023). For details, please refer to our response to the corresponding detailed comment.

Detailed comments:

Abstract: the result related to the lack of proper representation of the surface scattering layer over melting sea ice is missing from the abstract. I believe it is relevant to include it.

The reviewer points out an important issue, namely the need for improvements in the albedo parameterizations to account for an SSL. Although we have not quantified the impact of SSL on surface albedo, we have observed an underestimation of modeled surface albedo when only bare ice is considered in the model after snowmelt. We added the following statement in the abstract:

<u>"The lack of an adequate model representation of the surface scattering layer formed on bare ice contributed to the underestimation of surface albedo in summer."</u>

line 176: "…where sea ice is further divided into snow-covered ice (subscript s), bare ice (subscript bi), and melt ponds (subscript mp)": could you please add a comment on which ice category the "surface scattering layer (SSL)" (also called "white ice") belongs to? It is not snow but very much resembles it, being much more reflective than bare ice (for the definition of SSL see e.g. https://online.ucpress.edu/elementa/article/11/1/00103/195863/Evolution-of-the-microstructure-and-reflectance-of and https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2006JC003977).

From your Table 3, the surface scattering layer would belong to "snow-covered ice" category when looking at the albedo intervals.

The evolution of the surface scattering layer (SSL) on bare ice is not considered in HIRHAM-NAOSIM. We agree that the SSL is an important component for albedo parameterization. However, it has not yet been accounted for in HIRHAM-NAOSIM by a separate class. For the classification of surface types from camera observations, we combined SSL/white ice into a common class "snow-covered ice/white ice" due to similar reflectance properties.

We added the following:

"Note that the surface type "white ice", which results from a highly reflective scattering layer on top of melting ice (Macfarlane et al., 2023), is not explicitly considered in HIRHAM-NAOSIM. Due to its higher albedo compared to bare ice, white ice is added to the class of snow-covered ice in this work and classified accordingly based on camera observations during the measurement flights."

Fig 4: very nice Figure!!

Thanks.

Sect 3 and 4: reading these sections is extremely painful because of the unclear text and imprecise vocabulary. The logical rigor of the sentences is poor as there are often missing logical steps in the explanations. The text is not suitable for scientific publication and needs to be extensively rewritten. I provide here only some examples of poor sentences

We thank the reviewer for her/his detailed suggestions to improve the content and language.

line 300-301: "From that, we assume that the distribution shown for the modeled surface albedo is biased to higher values, since the cloud cover is overestimated." This is a quite badly expressed sentence and concept. Maybe you mean something like "Based on these results, we argue that the match between satellite- and model-derived surface albedo medians results from the compensation of two opposite model biases: the overestimation of modelled clouds, which caused a positive bias in modelled albedo, was compensated by a negative bias in modelled clear-sky albedo." Do you agree?

Yes, and you have understood what we were trying to say. We have improved and revised this text considerably to make the text clearer. The adjusted text reads:

"As the modeled surface albedo is cloud cover dependent, the representation of the clouds in the model must be taken into account to evaluate the modeled surface albedo. While the aircraft and satellite observations showed mostly cloudless conditions, the model calculated a cloud cover of about 100 % in most areas. Based on these results, it can be assumed that the match between satellite- and model-derived surface albedo medians results from the compensation of two opposite model biases: the overestimation of modeled cloud coverage, which caused a positive bias in modeled surface albedo, was compensated by a negative bias in modeled cloudless surface albedo."

lines 315-316. "At end of March, a distinct minimum of sea ice coverage (0.86) was simulated for the area covered by the flight on 3 April 2018 leading directly to the minimum of the surface albedo." Totally unclear sentence, I did not manage to guess what you mean.

We have revised this text to clarify what is meant. It reads now as follows:

"The temporal variation of the modeled surface albedo is illustrated in Fig. 5d. Each individual line represents the time series of the area-averaged surface albedo for one of the seven overflown areas. In addition, the mean measured surface albedo (including standard deviation) on the corresponding day is shown. […] The albedo time series of the area overflown on 3 April 2018, shows a pronounced albedo minimum in late March associated with a minimum sea ice cover (0.86).

line 316-318: "The corresponding measured areal-averaged surface albedo shows, on the one hand, a much greater spatial variability and, on the other hand, a clear tendency towards smaller surface albedo values. This tendency…" Please rephrase, and not use the word "tendency" if you are not showing a decreasing/increasing trend in your time series, it is very misleading. Do you mean that area-averaged modeled albedo is positively biased compared to area-averaged aircraft observations? If so, write it clearly.

We rephrased that part as follows:

"In general, the measured surface albedo shows much greater spatial variability but smaller averaged values than the model. This positive bias of modeled surface albedo cannot be explained by a lower sea ice coverage modeled with HIRHAM-NAOSIM."

lines 332-333: "This partly explains the difference in the distribution of modeled and measured surface albedo, in particular for the surveyed regions on September 8 and 13." You did not show this result: either you show the plot, or you remove this sentence.

Perhaps the word "distribution" is misleading here. From the colors in Fig. 6a, one can conclude that the modeled albedo for these two days has a negative bias. The average values are shown in Fig. 6b, which support the statement.

"This could partly explain the difference between the modeled and measured surface albedo. The modeled albedo map shows a negative bias compared to the measurements along the flight path (overlaying brighter points in Fig. 6a), especially for the flights on 8 and 13 September."

line 336: "Cloud effects are small, as mostly a full cloud coverage was modeled by HIRHAM-NAOSIM." This sentence cannot be understood if you don't include all the logical steps. Do you mean that "Clouds did not significantly affect the temporal variability of modeled and observed surface albedo because both observations and model simulations were carried out in overcast conditions"?

In general, greater variability in modeled surface albedo may also be caused by variable cloud conditions, since the albedo parameterization distinguishes between cloudy and cloudless conditions. For the MOSAiC-ACA cases presented here, however, both the model and the measurements show predominantly cloudy conditions. That's why we argue that clouds do not contribute to the modeled variability here.

We rephrased the sentence:

"Since HIRHAM-NAOSIM mostly simulated a cloud coverage of 100 %, the variability of the surface albedo cannot be due to the use of different parametrizations for cloudy and cloudless conditions."

line 337-338: "The measured areal-averaged surface albedo shows best agreement for the region overflown on September 2, although parts of the northernmost section of this flight path were underestimated by the model." It seems to me that part of the northernmost section of that flight is overestimated by the model (the western part) and part is underestimated (the eastern part).

Indeed, the flight section following mostly the same latitude (northernmost section) is showing partly an overestimation of the modeled albedo. We changed it:

<u>"The measured area-averaged surface albedo shows best agreement for the region overflown on September 2, although the surface albedo along the northernmost section of this flight path was partly overestimated by the model."</u>

Sect 3.3.3: it would be good to explain why you decided to use the data from one RAMSES spectral albedo station and not from other albedo stations (there were several broadband albedo stations and other RAMSES stations), especially because for this study the spectral to broadband albedo conversion. Wouldn't be more straightforward to apply broadband observations? How representative of the MOSAiC ice floe surface the data collected at the selected station are?

We agree, that all other data sets of albedo measurements made during MOSAiC have great potential. However, we are sticking with the RAMSES data, mainly because of its continuity. The 2020R12-RAMSES station was installed during leg 3 at the L3 distributed network site. It was operated between April 24, 2020 and August 07, 2020. We selected the RAMSES station because it is independent of the logistic gap during MOSAiC, where observations had to be interrupted. Since the time frame of this logistic gap exactly covers the transition from dry to wet snow during the onset of melting, we used the autonomous measurements for comparison with HIRHAM-NAOSIM to have a continuous data set available.
We are aware that the data from the surface flux stations also provide almost continuous time series of irradiance measurements during MOSAiC. The data sets have already been used in another publication by Foth et al. (2023), which is currently under discussion in TC. In this work, the HIRHAM-NAOSIM model was evaluated against the flux station data, focusing on the changes in the snow albedo parameterization with respect to clouds (presented in Jäkel et al., 2019). To avoid repetition, we did not use this data set in our study. We refer to this study in Sec. 1:

"<u>A comparison of the modeled surface albedo between the revised model and the earlier version was presented by Foth et al. (2023). They evaluated both model versions using measurements from two flux stations deployed during MOSAiC. They found that the revised snow surface albedo parameterization led to a more realistic simulation of surface albedo variability during the snowmelt period in late May and June."</u>

Foth, L., Dorn, W., Rinke, A., Jäkel, E., and Niehaus, H.: On the importance to consider the cloud dependence in parameterizing the albedo of snow on sea ice, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-634, 2023.

And in Sec. 3.3.3 we added:

"During MOSAiC, the seasonal evolution of the surface albedo was measured by autonomous radiometers. <u>In this study, data from one of the RAMSES stations (2020R12, following the notation of Tao et al., 2023) were used. 2020R12 was deployed on second year ice at site L3 of the MOSAiC Distributed Network (Nicolaus et al., 2022). This data set provides almost continuous time series of irradiance measurements between 24 April and 7 August 2020, which in particular allow the observation of the</u>

transition from dry to wet snow during the onset of melting. We applied the two corrections according to Eqs. (1) and (2) to the ground-based observation of the autonomous radiometers..."

As requested by the other reviewer, we have referred to the work of Light et al. (2022) and mentioned the role of autonomous measurements:

"Simulations and ground-based measurements of the seasonal evolution of surface albedo during MOSAiC were previously presented by \cite{Light_2022}. The authors used an Earth system model (1° spatial resolution) for comparison with surface albedo measurements manually made along three survey lines. These measurements could not be performed with the same high temporal frequency during the complete campaign for logistical reasons. Therefore, the transition from dry to wet snow during the onset of melting was less well observed than in our study, which relied on autonomous measurements from a radiation station."

line 359: the reference (Tao et al., 2023) is missing.

The manuscript by Tao et al. "Seasonality of spectral radiative fluxes and optical properties of Arctic sea ice during the spring-summer transition" is still in review (in Elementa). We will remove the citation in case the manuscript is not accepted in time.

Tao, R., Nicolaus, M., Katlein, C., Anhaus, P., Hoppmann, M., Spreen, G., Niehaus, H., Jäkel, E., Wendisch, W., and Haas, C.: Seasonality of spectral radiative fluxes and optical properties of Arctic sea ice during the spring-summer transition, submitted to Elementa: Science of the Anthropocene, 2023.

line 368-370: "On June 21 and June 22, both data sets showed a similar surface albedo, even though the spatial variation of the satellite product was smaller than the temporal variability of the ground-based surface albedo measurement within a day." In my opinion, comparing spatial variability of one data with temporal variability of another data is not meaningful. You need to better elaborate, otherwise the comparison between in situ and satellite data is meaningless.

Actually, we want to point out that the spatial variability of the observed surface albedo within the footprint of the HIRHAM grid at a fixed time can be smaller than the variability of the surface albedo on a single day. However, since this statement may be deficient, because small-scale changes cannot be resolved by the satellite observations, we have adjusted this sentence:

"On June 21 and 22, satellite and ground-based measurements showed a similar mean surface albedo of 0.69. For the observed cloudless conditions, Eq. (1) can be applied to correct the radiometer measurements."

line 391, eq. 10: the way in which the equation is presented is misleading: the difference between measured and modelled net irradiance does not depend on albedo alone but also on the incoming irradiance. Please correct.

We agree and generalized Eq. 10 as follows:

$$\Delta F_{\mathrm{net}} = F_{\mathrm{net,model}} - F_{\mathrm{net,meas}} \quad . \tag{10}$$

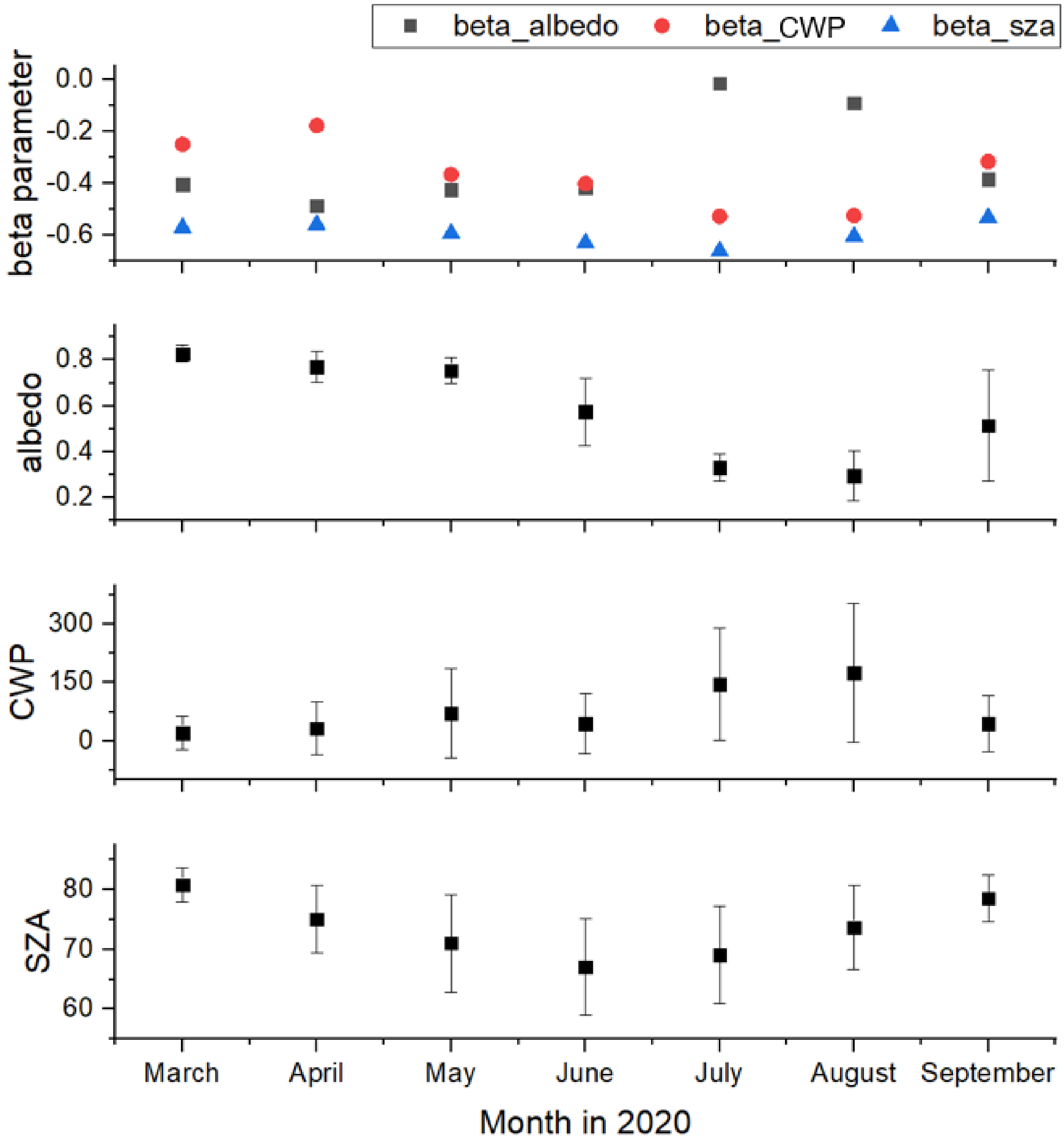
Sect 4.1, lines 406-417: the authors did a linear regression analysis to assess the relative impact of biases in albedo, solar zenith angle and modelled cloud water path on the bias in modelled net shortwave irradiance. I think that the results are very much dependent on the considered dataset

(March-April and September observations). A different dataset, with different spatial and temporal variability in albedo and cloud properties would provide different standard deviations with respect to model simulations, yielding a completely different result. For instance, I would expect that in summer, when albedo is lowest, cloud optical thickness is largest, and shortwave cloud radiative forcing is largest (most negative), the cloud std may cause a larger error in surface net shortwave irradiance than the std in albedo. Hence, I recommend considering the results from the perspective of the analyzed dataset and discuss the implication of different albedo and clouds conditions.

Yes, the reviewer is right, that the coefficients we derived in this study cannot be generalized and depend on the specific conditions of the selected data set. Therefore, we wrote: **"For the analyzed cases during MOSAiC-ACA and PAMARCMiP** we found the strongest impact of the surface albedo (β_α = -0.80), and less impact of the CWP (β_{CWP} = -0.38) and SZA (β_{SZA} = -0.23)."

Inspired by the reviewer, we have calculated the monthly beta-parameters for a specific area (flight area from September 2) as shown in the plot below. The three lowest panels depict the monthly means and monthly standard deviations of the surface albedo, CWP, and SZA. They support the reviewer's considerations that the beta parameters are variable over time and are associated with a reduction in albedo and an increase in CWP in summer. We further find that the magnitude of β_{SZA} in this time series is higher throughout the year than for the two time periods (April/March and September) we considered in the study. This is mainly due to the fact that we included complete daily cycles in the evaluation (all data points with SZA < 85°), whereas the data selection of PAMARCMiP and MOSAiC-ACA was limited to the times of the measurement flights. Since this is only a quick analysis, we will not make any quantified statements about potential changes of the beta parameters here. This could be the subject of a new study, but is beyond the scope of the current paper. Here, we have added the following:

"However, we expect a seasonal dependence of the standardized regression coefficients. According to Eq. (11), a stronger variability of the individual parameters contributes to a higher magnitude of β_j. In summer, clouds tend to have a high cloud water path with a high variability, while the surface albedo reaches its minimum. Therefore, it is assumed that the contribution of the surface albedo bias to the F_{net} uncertainty is reduced, whereas the model representation of cloud properties gets more relevant compared to the two periods shown in this study ."



Sect 4.2 and related text in Sect 5: often the expressions "surface albedo forcing" or "surface albedo effect" on the net shortwave irradiance are improperly used, as in reality you meant "impact of the surface albedo bias on the calculated net shortwave irradiance". This is very confusing. I recommend rewriting the text paying particular attention to the precision of the used vocabulary and espressions.

By analogy with cloud radiative forcing, which is defined as the difference between the net irradiance under cloudy and cloudless conditions (Ramanathan et al., 1989), we have used the term surface albedo forcing because it indicates the difference between the net irradiance derived for the modified condition (parameterized albedo) and the reference condition (measured albedo). However, we agree that this expression could be confusing, and therefore, we now omit this term as suggested by the reviewer. The text was changed accordingly:

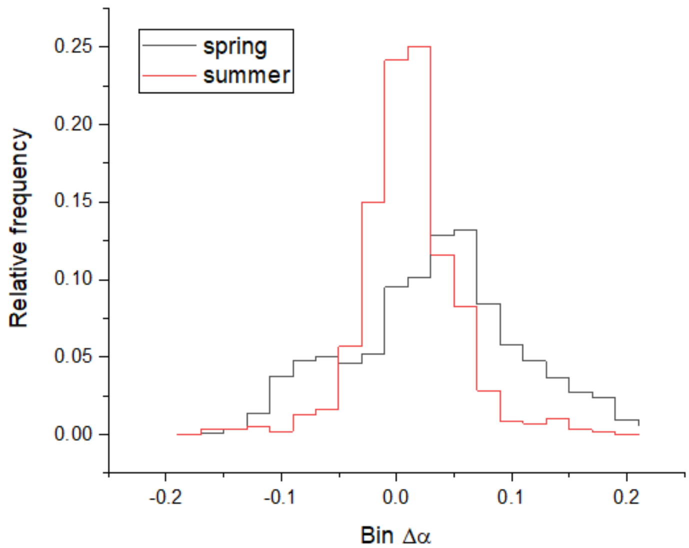
"<u>The maximum impact of the albedo bias on ΔFnet is derived for cloudless summer conditions (ΔFnet=±80 W m-2). For the same range of Δα in spring, ΔFnet is found to be less than half of its magnitude in summer (ΔFnet=±35 W m-2).</u>"

"<u>We investigated how the surface albedo model bias affects the balance between incoming and outgoing irradiance at the surface by calculating the net solar irradiance.</u>"

Ramanathan V. et al.: Cloud-Radiative Forcing and Climate: Results from the Earth Radiation Budget Experiment, Science 243, 57-63, DOI:10.1126/science.243.4887.57, 1989.

lines 445-446: "This indicates that a surface albedo bias in spring is less relevant for the absolute amount of the solar energy budget at surface than in summer.", and 495-497: "Since the maximum surface albedo effect on the net irradiance was derived for cloudless summer conditions, it can be concluded that the surface albedo bias is more relevant to the absolute amount to the solar energy budget in summer than in spring." Even if the surface albedo bias causes a larger bias in clear-sky surface net irradiance when the incoming irradiance il largest (in summer), it does not mean that it is more relevant for the summer than for the spring surface energy budget. In fact, cloudless skies are much more frequent in spring than in summer. However, the freezing temperatures make the albedo spatial and temporal variability and, thus, the bias in modelled albedo, much smaller in spring than in summer. I wish the authors could expand the discussion on this result, including references to previous studies.

Thanks for bringing this up. In contrast to the statement of the reviewer ("… and, thus, the bias in modelled albedo, much smaller in spring than in summer. ") we observed an increased albedo bias with a wider distribution in spring than in summer (see figure below). This may not be true for other models where the cloud dependence of the surface albedo parameterization is not considered. In our study, however, we see an overestimation of the modeled surface albedo, as the cloud dependence in the albedo parameterization is deficient for optically thin clouds, which contributes to the albedo bias in spring.



We rephrased the text in Sec. 4.2 and removed the sentence in Sec. 5 to weaken the statement that the surface albedo bias is more relevant in summer:

"<u>For the same range of Δα in spring, ΔFnet is found to be less than half of its magnitude in summer (ΔFnet=±35 W m-2). In spring, however, we observed from the flight measurements an increased albedo bias with a wider distribution (Δα=0.02±0.07) than in summer</u> <u>(Δα=0.00±0.04). This means</u>

that greater effects on the solar radiation balance between solar incoming and outgoing irradiance due to the surface albedo model bias can be observed in summer, but these are less likely than in spring."

The reason for the poor representation of the surface albedo by the model in spring was investigated in more detail with the help of Fig. 9c.

lines 472-473: "We conclude that a functional dependence, rather than a pure discrimination between cloudy and cloudless conditions, is required to properly describe the cloud effect on surface albedo." The advocated physical dependence of the broadband albedo parameterization on cloud properties (optical thickness) is much less physically consistent than the waveband-dependent albedo parameterization would be. Only a waveband-dependent albedo parameterization that at least distinguishes between visible and infrared regions can account for the cloud impact on albedo in a manner that retains the coupling and dependencies between the physical variables. Could this solution be applied in HIRHAM-NAOSIM? I invite the authors to consider this solution or at least to comment on it.

We completely agree that the use of a waveband-dependent albedo parameterization would be physically more consistent. However, the implementation of such a complex parameterization is very time-consuming and much more complicated than introducing a simple cloud dependence in the present broadband albedo parameterization. Furthermore, and indeed the simulation results indicate that the cloud effect on surface albedo can be reasonably reproduced using a broadband albedo parameterization with cloud dependence. Based on our results we argue that the problems in the present albedo parameterization are rather related to deficiencies in the simulation of clouds and the surface fractions than to poorly reproduced cloud effects. These deficiencies will not be remedied by switching to a waveband-dependent albedo parameterization. Therefore, we are convinced that our results using a simpler broadband albedo parameterization are helpful and indicate that we should focus on improving the representation of clouds and surface type fractions in the models next, before implementing the more complex albedo parameterization. Basically, we agree that a waveband-dependent albedo parameterization should be implemented on a longer perspective.

We added:

"In the absence of a waveband-dependent albedo parameterization, the consideration of a simple cloud dependence in the broadband albedo parameterization is able to mimic the cloud effect on surface albedo reasonably. The cloud effect might be further improved by a more sophisticated functional dependence on cloud cover or cloud water content, rather than a pure distinction between cloudy and cloudless conditions."

lines 488-489: "In particular, the surface albedo of the scattering layer classified as bare ice seemed to be underestimated." This is a critical issue: currently, in all sea ice schemes that I am aware of, the albedo of ice without snow is simulated as the albedo of bare ice, which is much lower than the albedo of the surface scattering layer. The surface scattering layer is completely ignored in the sea ice surface schemes, with the consequences that you have illustrated. I invite the authors to expand on this issue: what is the impact of ignoring the surface scattering layer on the surface energy budget? Could the surface scattering layer be modelled? Please refer to Macfarlane et al: https://online.ucpress.edu/elementa/article/11/1/00103/195863/Evolution-of-the-microstructure-and-reflectance-of

We followed the Reviewer's suggestion and added the following in Section 5:

"This is due to the fact that the model assumes bare ice instead of a surface scattering layer, which emerges at the top of the melting sea ice after the snow has melted. The SSL is a porous, granular, and highly fragile pillared structure on top of the ice, which effectively backscatters solar radiation and keeps the surface albedo of melting ice relatively high (Macfarlane et al. 2023). Due to the small-scale characteristics of the SSL, it is pretty difficult to relate the surface albedo of the SSL to the available variables of a climate model with spatial scales in the order of several kilometers. Consequently, the surface albedo of the SSL is a critical issue in the albedo parametrization. Since the albedo of bare ice is generally lower than the albedo of the SSL, the surplus of radiation energy at the ice surface may lead to an amplified melting of sea ice in the model."

Further, we rephrased the following (Section 3):

"We also assume that the predominantly modeled bare ice fraction with its low surface albedo contributes to the model bias. In the field, however, the surface albedo of the melting ice remained relatively high due to the presence of a brighter SSL, which is not taken into account in HIRHAM-NAOSIM."