Review of: ”Insights into the low-level single-layer stratiform cloud optical depth feedback based on two decades of observations at the North Slope of Alaska.” by Coulbury et al.

General remarks

The manuscript analyzes observations of cloud properties by the Atmospheric Emitted Radiance Interferometer (AERI) at Utqiagvik, North Slope of Alaska. The data base extends to about 20 years which allows the authors to investigate long term trends in cloud properties. Different aspects are covered by the analysis ranging from simple trends in cloud optical thickness, changes of the cloud radiative effect, the cloud feedback to surface temperature and identifying controlling factors for cloud property changes.

Although limited to a single specific Arctic location, in general, an improvement of our understanding of these cloud processes by observational constrains would be valuable for arctic climate modelling. However, the study lacks of significant issues that I will explain in detail below. Most evident is that most results are not statistical significant. Still the authors attempt to interpret the calculated numbers. These challenges in interpretation are increased by the selection bias introduced by the AERI retrieval and not considering the changing ocean/sea ice conditions at Utqiagvik. In combination with the low confidence of the results, I doubt that any robust conclusions can be drawn from this study. I also do not see, how this can be improved by simple changes of the manuscript.

Therefore, I have to reject the manuscript. I do not see a way how to adjust the study to make it reasonable. The authors first have to investigate how their measurement approach affects the results. But still it is questionable, if data from a single station/location and using only a 20 year time series can show significant trends and correlations. Maybe including other Arctic stations (still only a limited number and often in specific local environments) might help. But this is not easy to achieve. Another option might be to resubmit a new manuscript that focus on one of the aspects of Sections 3.1-3.6 and elaborates this in detail. E.g. showing a 20 year data set of CRE without trend analysis or correlations would alone be valuable. The seasonal cycle still could be studied in a more extensive way than in current Fig. 3. The authors could also explore how CRE estimates from different instruments data sets (many are available at NSA, radar, MWR,...) compare. This would address also the limitations and potential selection bias of the AERI based CREs. These are just some ideas to reduce the frustration caused by my review. But the authors may have their own ideas.

Major comments

Significance of results

Most detected trends are week and statistically uncertain. E.g. in Fig. 4 the confidence interval for all results are at least a factor of two larger than the magnitude of the detected feedback. To me, this means, that not even the sign (positive and negative feedback) can be estimated with sufficient certainty. With such high uncertainties, nothing can be concluded from the results. E.g.: L321 - 324 can not be stated.

The values of CRE in Fig. 3 are very low compared to literature of cloud radiative effects (e.g., Shupe and Intrieri, 2004). Is this due to the filtering for thin low clouds in the data set? Then, what do we learn, if only clouds with low CRE are analysed? Additionally, the trends presented in Fig. 4 are of similar magnitude than the climatological CRE in Fig. 3. Does this mean, a 1 K temperature change causes the CRE to almost double? This seems to be unrealistic given that we already observed a strong increase of Arctic temperatures.

In L328-335 the authors confirm, that their results are not significant. If this is the case, the authors should not even try to interpret their results. The authors also add, that their analysis can not be considered as long-term feedbacks doe to the limited time period (L333). Here the authors also contradict themselves in their discussion:
”This indicates that much of the variance.... can be explained by factors other than globally averaged surface temperature.” is followed by ”...it is important to consider how cloud changes, as mediated by global surface temperature trends, are impacting the radiative budget at the Arctic surface”.
If temperature does not explain cloud changes, then there are no cloud changes as mediated by global temperature trends.

Limitation and impact of cloud type

The significance of the results is further reduced by the limitations of the AERI/MIXCRA retrieval of cloud properties. As mentioned by the authors, the retrieval will converge only for low, optically thin, single-layer clouds. Thus, only a subsection of clouds are studied. This makes interpretation challenging or even impossible, because observed trends can be due to a trend of cloud properties but also due to a trend in the retrieval performance. Two issues:
I) The data is automatically limited to $COD < 50$. What if a climate trend pushes a cloud with $COD = 49$ into $COD = 51$? Then this cloud is removed from you analysis.
II) Similar the convergence of MIXCRA might change with increasing temperatures or changing cloud properties. What if this will mostly affect thicker clouds and less thin clouds. This also might induce an artificial signal in your trend analysis. As illustrated in Fig. 1 and discussed

by the authors in L300, there is a seasonality in the retrieval performance. What if there is also a climate trend of the retrieval performance over 20 years? To me, this makes it challenging to interpret any result of the analysis. It might simply be biased by the retrieval and the selection of a subrange of clouds. It makes me also doubt if Fig. 5-7 can be used to derive any conclusions.

Of course these are all limitation of the approach presented in the manuscript. If the trend signal would be higher than the uncertainties in estimating cloud properties, CRE, etc. this would be OK. But given the fact, that already the trends are not statistically relevant, the uncertainty introduced by the observations and retrieval limitations will make any conclusion questionable.

All this makes me wonder, why the study is limited to AERI measurements? Why not all data from the NSA are used? Observations at Utqiagvik offer a whole range of cloud and atmosphere remote sensing. Another approach would be to measure CRE directly independent to any retrieval algorithm. A recent study by Lubin et al. (2026) used microwave radiometer and broadband solar and thermal-IR fluxes from pyranometer and pyrgeometer to analyse trends in cloud properties at Utqiagvik. It seem necessary to argue, why the authors did not include these independent data in their analysis. The results of this manuscript also needs to be contextualize with respect to Lubin et al. (2026).

Cloud controling factors

I don't agree, that the authors can detect cloud controlling factors. The authors simple look at correlations by a multi-linear-regression. A correlation does not tell anything about causes and consequences. E.g. higher T_{850} may cause a higher COD, but it can also be that a higher COD causes a higher T_{850}. The strong correlation of cloud optical thickness and OCPHILIC might not be related to aerosol-cloud interactions. It is known, that the production of organic particles is less in presence of sea ice. What if less sea ice leads to more clouds but also to more organics? This also will provide a nice correlation but OCPHILIC does not control the clouds. And the same holds for most other parameters/correlations.

The authors also miss to address the uncertainties of all measured quantities used in Section 3.5. and how this will affect the derived correlations. Significance is only quantified by confidence of the regression. And often this significance is limited. The authors state themselves: L439: "this impact could....", L440: "This could...", L441 "... but a clear analysis of this relationship is precluded by a low correlation and large error." or L443: "highly uncertain". Reads for me like a lot of speculation and no clear conclusions.

I also wonder why the surface albeo/sea ice cover is not considered in the analysis as a major parameter characterizing Arctic environments. The impact of sea ice on cloud properties is only briefly discussed, almost neglected, in the manuscript. This starts with L50-51, where only thermodynamic parameter are considered to change cloud properties. I would doubt, that this fully holds for Arctic environments. Sea ice is a major component that regulates the conditions at the surface. E.g., sea ice cover changes the water vapor supply (latent heat flux) which is a major control of boundary layer clouds (e.g., Saavedra Garfias et al. 2023). And the location of Utqiagvik is prone to seasonal changes of sea ice cover. This is recently shown by a study of Saavedra Garfias and Kalesse-Los (2025) for surface-coupled clouds observed at Utqiagvik. This may also explain some of the results. E.g., **L7-8**: Why the COD changes are unsymmetric in season? Could this be linked to sea ice? Spring and Autumn experience a more significant change in sea ice cover than summer and winter. That's why I thinks, that Section 3.5 can not be presented without including sea ice cover in the analysis and reasoning.

Cloud radiative effect

The authors approach to calculate CRE following Zhou et al. (2022) might introduce significant uncertainties. In Zhou et al. (2022) I don't see an evaluation of the method for high latitudes. How confident are you, that your application to NSA is providing reliable CRE?

The approach by Zhou et al. (2022) is build for a fast evaluation of climate models. In this study a 20 year time series at a single location is used. For a single location I assume that it is possible to apply more precise methods to calculate CRE, e.g. run a radiative transfer code tailored for your location and using all observations that are available. NSA provides 6-hourly radiosonde date, cloud base from ceilometer,.... Everything to run a radiative transfer model is available. What is the benefit of using Zhou et al. (2022) instead? What uncertainties do the authors expect from this method?

The manuscript also does not reveal how changes of the local surface albedo in Utqiagvik are treated in the calculation of CRE. The solar component of the CRE strongly depends on surface albedo. As surface albedo (snow, sea ice, ocean) changes a lot at Utqiagvik (seasonally and likely also climate trends are obvious), this needs to be considered.

Linear trend

In Fig. 9 I simply can not identify such a strong trend (red line) by eye. Especially when extreme month are weighted less, the trend should not exceed the majority of data points. It might help to show the unfiltered and filtered time series to convince the readers. But I still would doubt, that a statistically significant trend will become visible. If this can not be convincingly shown for the COD time series. I would also question if trends in Section 3.3 can be identified.

Could it be, that the problem results from using monthly data in combination with your filtering?

Fig. B1 also does not look convincing. The smoothed CF anomaly time series does not represent the full time series in my view. Any issue with the fourier transformation? However, if the authors want to go for a 21 year trend, annual averages should be sufficient. Then you would not struggle with filtering monthly variability.

List of specific comments

Abstract L10: Are these parameter measured data or reanalysis data? Mention the source of the data in your abstract.
Abstract L11-12: Is this the same conclusion as stated in L3-4 ? Are the conclusion given above also significant and from the same long data set? I suggest to restructure the abstract.
L82: Section title: R_{eff} is also retrieved via MIXCRA.
L86- L254: I would emphasize that this effective radius represents the entire cloud layer. You do not retrieve a height resolved profile of R_{eff}. Thus your R_{eff} represents the entire cloud. Folloe up question: Is this definition of R_{eff} considered in estimating LWP/IWP? What profiles are assumed for the cloud particle sizes?
In this respect, Eq. 5 refers to homogeneous cloud layers, neglecting the increase of particle size and LWC with altitude. This might be OK, when your assumption on the cloud verticle profile in the MIXCRA retrieval is the same. Then your R_{eff} corresponds to a homogeneous cloud with droplets of this size. If you have different assumptions on the cloud profile, this could result into biases.
L93: LBLDIS: explain acronym
Section 2.2: The MIXCRA description is quite lengthly and detailed. On the other hand, the motivation of the manuscript puts the focus on cloud feedback analysis not on retrieval description. The analysis done in the manuscript also does focus on climate analysis and not on the evaluation of the MIXCRA results. Although it is important, that measurement uncertainties and limitaions are provided, I don't thinks such a detailed MIXCRA desciption is not needed and available publication can be cited here.
L111: "dual phase". The most common phrase is "mixed-phase cloud". Or does "dual phase" mean something different?
L119: What is RU?
L167: This equation could be repeated and included here.
L172: With respect to the cloud radiative effect, averaging COT might cause trouble because the relation between e.g. transmitted radiation and COT is not linear. Averaging COT likely overestimates the total cloud radiative effect, especially for low COT. (See your statement in L175-176)

L205: Your trend analysis in based on only 20 years of data. You tried to remove internal variability, especially an oscillation of 9 year period was removed. You argue that this mit be related to PDO. Can you prove this? Otherwise you might also remove some of the natural trend. A 9 year period only fits twice into your entire time series. Depending on the phase of this oscillation, it might significantly contribute to the total trend, when not removed. A non-perfect or over-emphasised filtering might still affect your analysis. Can this be quantified?
It might be my impression, because the applied methods do not fall into my major expertise. But to make your approach trustworthy, I strongly suggest to add some illustration of your method by example plots (E.g., the histograms discussed along section 2.3.1. or the regressions discussed in L250.)
L206: What is PDO?
L213 and Eq. 3: Cloud fraction was written as CF before. It is somehow misleading when the anomaly of cloud fraction is represented by ΔC and not ΔCF.
Eq. 3: Are the quantities ΔCRE and the kernel $\delta R/\delta C$ also binned? If yes, I would suggest to add the dependencies, e.g., $\Delta CRE(COD, CBP)$. Otherwise it looks as if your CRE anomaly only depends on cloud fraction anomaly. This seems to be too simple.
L221-224: This reads as if you analyzed MODIS data in your study. I guess, this only refers to the Tan et al. (2019) study.
L288: Can you mention the method Zhang et al. (2019) used to derive cloud properties. We know, that loud properties can differ depending on the retrieval the method (active/passive remote sensing...).
L307: add "...by the cloud"... also the surface emits and reflects.
Fig. 1: "Climatology" what does it mean? Are "monthly averages" plotted? If yes, use the more specific name.
Fig. 2: Does the plot show total numbers? A relative fraction is also interesting. Or add the number of failed retrievals.
Fig. 5: How annual values can differ from winter values, when low temperatures were only observed during winter?
Fig. 7: Y-axis labels are missing.
Fig. 8: Why this analysis was not split into season?
Section 3.6: Why authors started with the most challenging analysis of CRE and cloud feedback and end with the most simple analysis trend in optical thickness? I suggest to sort the manuscript by starting with simple analysis.
L530: I could not find the data set. Doi ist not working. The data set also need to be included in the bibliography.
L700: Tan et al. (2024) is listed twice in your bibliography.

References used in the review

Lubin, D., Zou, X., Mülmenstädt, J., Vogelmann, A., Cadeddu, M., and Zhang, D.: Surface radiation trends at North Slope of Alaska influenced by large-scale circulation and atmospheric rivers, Atmos. Chem. Phys., 26, 295–311, https://doi.org/10.5194/acp-26-295-2026, 2026.

Saavedra Garfias, P., Kalesse-Los, H., von Albedyll, L., Griesche, H., and Spreen, G.: Asymmetries in cloud microphysical properties ascribed to sea ice leads via water vapour transport in the central Arctic, Atmos. Chem. Phys., 23, 14521–14546, https://doi.org/10.5194/acp-23-14521-2023, 2023.

Saavedra Garfias, P. and Kalesse-Los, H.: Observed modulation of wintertime Western Arctic mixed-phase cloud properties by sea ice conditions, their long-term variabilities and trends, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-2327, 2025

Shupe, M. D., and J. M. Intrieri, 2004: 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.