We would like to thank the editor for the opportunity to publish our manuscript at EGUsphere and to the nine referees for their thorough evaluations with constructive comments that will improve the manuscript greatly. In the following, we will address the referees' comments point by point. We mark the comments given by the referee in red, provide our answers and comments in black, and indicate how we will address the amendments in the manuscript that we plan to submit upon editor's decision in green. Note, that we add the concrete amendments planned not at all points as this in parts make the replies less readable and only where we will add/adapt a section directly and concisely. In many parts we only conceptually explain how we will adapt content, while the exact implementation remains until the revised version.
We would like to announce that should the manuscript advance to the revision stage; Verena Haring from the Department of Biology at the University of Graz will be included as a co-author in recognition of her valuable assistance with the point-by-point responses.

Tiago Silva, on behalf of all co-authors.

RC1: 'Comment on egusphere-2024-2571', Rúna Magnússon, 20 Oct 2024

Dear authors, dear editor
Thank you for inviting me to review "Bio-climatic factors drive spectral vegetation changes in Greenland", by T. Silva et al. for Biogeosciences. The manuscript explores the role of a wide range of bio-climatic factors in explaining satellite-derived vegetation dynamics in Greenland. The authors aimed to identify which sub-surface and above-surface climate factors were associated with greening in Greenland, and how such associations differed among ecoregions and latitudinal and altitudinal gradients. They report that increases in the duration of the thermal growing season show the strongest association with greening, with additional influences of snowpack dynamics, and differential strength of association across regions and altitudinal and latitudinal gradients.
This study is relevant in the context of rapid ongoing climate change in the Arctic, observed dynamics of "Arctic greening" and their implications for the future functioning of tundra ecosystems. The authors analyze a substantial amount of data for a large region, using various sources and environmental disciplines in a holistic and, generally, appropriate way. The manuscript is well within the scope of Biogeosciences and presents a relevant and timely case. I do, however, have several major concerns about some of the methodological choices and the structuring and argumentation of the work. I advise a round of thorough revision and rewriting of substantial parts of the manuscript before it can be considered for publication. This has resulted in a rather lengthy review report, but I would also like to stress that many of the points I raise are interrelated or specific examples of the major points, so I hope it is not discouraging. I am sure the ms will find a good home in a respected journal.
I have performed this review together with 7 MSc students for an open review course assignment at Wageningen University. Their help has been valuable, and they appreciated the opportunity to learn from this ambitious and relevant paper, and to contribute to the scientific publishing process. We have all enjoyed this activity and

we wish you all the best as the manuscript comes to full maturity!

Rúna Magnússon,
with input from Annika Robben, Djordy Potappel, Aron den Exter, Muriël de Vries, Rikuto Shinagawa, Yente Reniers and Yorick Kwakkel.

Major comments

1. I hope the authors can make clarify how the potential mismatches between AVHRR and VIIRS NDVI products (e.g. masking differences) have been accounted for during statistical analysis and trend detection. Explanations on how this was done are sparse and not sufficiently clear to understand the implications.
Beside adding the shaded min-max range in Fig. 2 (that I also don't fully understand the procedure behind, can this be clarified?),
how did you prevent the use of two different records and sensors from affecting your temporal trends? And especially, how do you prevent this from unduly influencing the comparison between 2008-2023 and 1991-2007, that you describe in L. 380-392? This appears to be based on counts of NDVI > 0.15, where differences in bandwidth and snow/water/cloud detection easily become problematic. Miura et al. (2012) may be an appropriate source to evaluate the validity of trend detection across two satellite platforms, and you may want to statistically test for absence of trend breaks coinciding with the switch from one platform to another.
The NOAA Climate Data Record (CDR) of AVHRR NDVI - Version 5 and the NOAA CDR of VIIRS NDVI - Version 1 are developed by Eric Vermote and colleagues (Vermote et al. 2018 and 2022) for NOAA's CDR Program. Both records have been processed considering the same atmospheric features as in Miura et al. (2012) and both processed records are posterior to Miura et al. (2012) proposed correction. Also, the correction proposed by Miura et al. (2012) is not assessed in polar regions, which may contribute to additional uncertainties.
We follow similar approaches of recent literature (e.g., Madson et al. 2023, Pourmohamad et al. 2024) that make use of the full AVHRR NDVI and VIIRS NDVI without additional corrections. As stated in Section 2.2, Vermote and colleagues for NOAA's CDR Program use MODIS to spectrally calibrate AVHRR (Vermote et al., 2018) and VIIRS (Skakun et al., 2018). As NOAA does not provide an overlapping period for AVHRR and VIIRS, we are unable to compare both processed products and quantify biases in polar regions. Nevertheless, we will make sure that we add to the discussion that the potential mismatches between AVHRR and VIIRS NDVI products cannot be discarded and, in a revised version we will provide the greenness trends before and after the sensor change in order to assess potential mismatches between sensors, bearing in mind that differences can also rise from other sources such as the interannual variability of the atmospheric conditions before and after 2014. In addition, we address the variability range during the AVHRR period in Figure 2 and discuss it in detail under a minor comment (point 9) below.

2. Your methodology is ambitious and extensive, which is laudable. It does however lead to many choices during the processing of the data, and not all of these have been properly backed or described yet. Examples include (1) the use of 0.15 as an NDVI

threshold without a reference, (2) the described use of CryoClim data that only go to 2015 without any visible inclusion of these data throughout later analyses, (3) why has only altitude, and not for example slope aspect, been included into the study? These, and further examples, are given in the minor comments. I suggest that the authors critically go through every step in the methodology and check whether all choices are described in sufficient detail for an independent reader to reproduce the study, and that choices are back-up either by literature, data or statistics. If needed, details on processing can be described in a supplementary methods section to prevent disruption of the flow of the main text.
Thanks for your kind work endorsing our ambitions!
Indeed, we will carefully review the clarity of the Methods to assure reproducibility in future interdisciplinary studies. Some specific comments below:
(1) Liu et al. (2024) is cited in the first sentence of Section 3.1, which shows sparse vegetation, with NDVI of approximately 0.15 at the start of the growing season at Disko Island. Additional online source is provided in a minor comment (point 7) below.
(2) long-term reanalysis products commonly use several datasets that do not cover the full period of the reanalysis. Additional examples are provided in a minor comment (point 5) below.
(3) for a matter of simplicity, the investigation of vegetation distribution as a function of slope and aspect was not considered so far. We will follow the advice and will include similar charts as for Figure 5 on surface slope and aspect in the supplementary material, whenever needed to support our Results. More information is provided in a minor comment (point 40) below.

3. From L. 227 onwards, it reads as if the distinction between methods, results and interpretation of results (discussion) is lost. For example: results and maps are presented in the methods in L. 227-247. New information on choices of processing, variable selection and statistical tests (Pearson correlations) are introduced in the results in L. 275-305, L. 311-314 L. 380-384 and many other places. Throughout the entire (lengthy) results section, interpretations are added that go beyond the statistical results of your own methods. Lines 320 -350 for instance are very speculative for a results section, and other paragraphs and show similar interpretation or speculation. These would be better suited for the discussion and require backing by references. Please rewrite the methods-results-discussion in such a way that: (1) all methodological choices and tests are explained in the methods (2) only numerical and statistical outcomes are presented in results (with a minimum interpretation to make the results understandable, e.g. writing out abbreviations and description of patterns) and (3) interpretation and relation to unmeasured mechanisms such as permafrost, latent heat processes or photosynthesis are only kept for the discussion.
Thank you for pointing this out. We will revise the Methods, Results and Discussion to make the differentiation between these sections more evident.

4. In the results section and abstract, observed greenness dynamics are attributed to processes such as nutrient dynamics and permafrost. This gives the reader the impression that such variables were included or that you can at least confidently

attribute greening dynamics to such processes. Given the set of bioclimatic factors that were included, however, I doubt whether you can make such claims. These processes can be touched upon in the discussion, with support from literature, but should not be presented in a way that readers might think that these are actual conclusions from this study. I also think that to properly discuss their role in the discussion, you will need to evaluate several lines of reasoning more critically: are the subsurface products (soil water and soil ice) that you include, given the limited representation of subsurface dynamics in the used reanalysis products, actually representative of permafrost conditions or hydrology? How can you better argue the role of snowmelt rates in relation to microbial activity and nutrient dynamics, especially to an audience that may not be familiar with works such as musselman et al.? Because at first it is very counterintuitive that shallower snowpacks melt more slowly and with the current explanation provided, this line of argumentation is very hard to follow. I suggest you evaluate to what extent your bioclimatic variables are representative of processes such as permafrost dynamics and melt rates and nutrient dynamics, discuss their potential roles in the discussion section, and refrain from making any hard statements about their role in the abstract/results/conclusion sections.

The Copernicus Arctic Regional Reanalysis (CARRA) -- Full system documentation (Schyberg et al., 2020) reports that areas with permafrost in Greenland are not fully described in the present model version. Nevertheless, the HARMONIE-AROME regional numerical weather prediction model (Bengtsson et al., 2017), uses SURFEX (Masson et al., 2013) a multi-layer surface model that accounts the interaction between soil-biosphere-atmosphere, computing specific models dependent on the surface type (e.g., vegetation, soil, snow), allowing soil water phase changes and enabling runoff over frozen and unfrozen soil. This implies that even though the reanalysis utilized does not integrate a surface model incorporating permafrost or nutrient dynamics, there is still an accounting for energy and mass exchange, achieved through different modelled processes that are influenced by the surface type.

Given the limited access to nutrients in the High and Low Arctic, organic matter decomposition is inferred due to microbial activity. The water availability due to permafrost and snow melt could stimulate decomposition of organic soil. We will discuss and review further in-situ studies supporting our interpretation of the modelled results.

We acknowledge that our reference to the work by Musselman et al. (2017) has not been clear enough. In the revised version we will add: Due to rising temperatures, snow begins melting earlier in the year. This results in the maximum snow depth being reached earlier than in the past. However, early in the year, the available energy from the atmosphere is still relatively low, which means the snow may melt slower than it used to be before the warming. This slow melting is especially noticeable in thinner layers of snow, which need less energy to start melting. On the other hand, thicker layers of snow persist until late in the year when conditions like higher sun angles provide more energy, leading to a quicker and more intense melting process.

Minor comments

1. The writing could be improved by splitting up some very long compound sentences into shorter ones. I provide some examples in the "technicalities", but I recommend a thorough re-reading for writing style and grammar.

Thank you for this very relevant point. We will thoroughly review and revise our sentence construction in order to improve readability and understanding.

2. The abstract ends with the conclusion that you "identify a set of bioclimatic variables" and that you provide a "basis to validate bioclimatic indicators from climate models". Your conclusions section states more or less the same. I suggest that you reflect more specifically on how exactly your findings help to achieve this (more to the point). This will hopefully also better explain how you advance the field, since the role of growing season onset and snowmelt timing are already well established in Arctic ecological studies.

Thank you for pointing out that we need to further develop our conclusions to show how our work advances the field. In our study, we show the potential of a polar-adapted, high-spatial resolution reanalysis product on capturing the combined role of a set of bio-climatic indicators with spectral greenness over a period of more than 30 years across the Greenland scale. We report the potential of a large-scale product for biogeographic research and discuss the application of a set of bioclimatic indicators in the validation of historical data (1991-2023) from global climate models that aim to capture future vegetation dynamics. We will make the core findings clearer and more concise in the conclusion of a revised manuscript.

3. 35 – 43. Several references seem out of place in this paragraph. I suspect you mean Bjorkman et al. (2018) instead of Metcalfe et al. (2018), since Metcalfe et al. (2018) does not deal with the type of findings you describe at all, and Anne Bjorkman's paper does. Sturm et al. (2001) is a rather old and case-specific (albeit popular) reference for shrubification of the Arctic. ITEX papers (e.g. Elmendorf et al., 2012) or syntheses (e.g. Mekonnen et al., 2021; Martin et al., 2017; Myers-Smith et al., 2011) would be more appropriate.

Thank you for identifying misplacement of references by the reference management. This will be corrected in the revised version. We also appreciate the suggestion of more up-to-date references regarding shrubification.

4. 66 – 68, this proposed increase in nutrient availability under deeper snow is at odds with your statements in the abstract, results and discussion that shallower snowpacks should melt more slowly. It should be clear from the introduction onwards which snowpack properties can be expected to facilitate faster or slower melt, and how would relate to nutrient cycling. If the literature on the influence of snow dynamics on microbial turnover and nutrient availability is ambiguous in itself, then I would refrain from making any statements about nutrients as a mediating effect between snow dynamics and greenness.

Increased snow during the cold season could allow more vegetation growth in the following warm season, as more snow provides insulation, less frost damage and an increase in water availability. However, the rate of snowmelt is essential for efficient meltwater percolation. This is particularly important in tundra regions, where soils are dry due to high drainage or low precipitation. In contrast, rapid melt will saturate the soil surface layer and run off. The interpretation of our results is that a considerable part of the dense/greenest vegetation is correlated with relatively shallow snowpacks and slow melt rates. This is particularly true in the southern ecoregions, where deep snowpacks are at higher elevations. Due to the decreasing trends in snow water equivalent (SWE_MAX) in the southern ecoregions, snowmelt starts earlier in the year, as they require less energy to melt, and the melt rates are also becoming slower. Ecoregion 2, which already has shallow snowpacks and does not show substantial trends in SWE_MAX, also displays a pattern of decreased snowmelt rates. We will expand on this complexity in a revised manuscript.

5. 111, here you mention the use of CryoClim data, that was chose to represent daily snow cover rather than the CARRA dataset. I do not see how this could be done since the data only goes to 2015, and this data product is not mentioned anywhere anymore in the remainder of the ms. Did you actually use it and if so, how? Perhaps it is a nice addition to incorporate data sources directly into Table 1 to resolve unclarities like this.
The assimilation of CryoClim data (ending in 2015) and, for example, MODIS (starting in early 2000s) on CARRA (from 1990 to present) is described in CARRA's Full System Documentation (Schyberg et al., 2020). The same often happens for automatic weather stations in Greenland. Although CryoClim has not been available since 2015, van der Schot et al. (2024) reports how CARRA performs against in situ measurements until 2023 across Greenland. Accuracy metrics are going to be provided as suggested in point 10.

6. 128-153: Can you give an indication of the match between AVHRRR and VIIRS? Calibration against MODIS does not seem to be the most relevant thing to mention here, since you do not use MODIS. See Miura et al. (2012), there seem to be some structural NIR differences and non-linear NDVI relationships between VIIRS and AVHRR?
There is no objective indication of the match between AVHRR NDVI and VIIRS NDVI computed by NOAA. AVHRR NDVI and VIIRS NDVI technical reports from NOAA as well as Miura et al. (2013), state different NIR and R bandwidths, additional to different algorithm corrections. As shown in Figure 2, AVHRR NDVI is generally lower than VIIRS NDVI. We attribute these differences in the first part of the time-series to less favourable environmental conditions for vegetation greening and NDVI retrievals. The linear relationships proposed by Miura et al. (2012) to correct AVHRR/2 ($y = 0.0412396 + 0.939953x$) and AVHRR/3 ($y = 0.0515123 + 0.872332x$), where y is VIIRS NDVI and x is AVHRR NDVI, would decrease the "corrected AVHRR NDVI" signal, consequently leading to higher differences between AVHRR and VIIRS NDVI. In order to get an indication whether interannual changes in green vegetation extent (more information about green vegetation extent in point 8) during VIIRS period are significantly different from the AVHRR period, we will statistically test the interannual changes in green vegetation extent for both periods.

7. 143, why did you use an NDVI threshold of specifically 0.15?

In contrast to other studies that either use the entire NDVI range or simply the maximum NDVI, we briefly explained in Section 3.1 that we only consider the NDVI that represents spectral greenness (NDVI > 0.15). While trends using the full NDVI range are heavily influenced by drying (NDVI transitioning from negative to positive values), trends using maximum NDVI in the high latitudes are likely to be influenced by temporal sampling artefacts (Myers-Smith et al. 2020).

In addition to Liu et al. (2024), who report NDVI at around 0.15 at the start of the growing season on Disko Island (in central West Greenland), we find that the United States Geological Survey states that "Areas of barren rock, sand, or snow usually show very low NDVI values (for example, 0.1 or less). Sparse vegetation such as shrubs and grasslands or senescing crops may result in moderate NDVI values (approximately 0.2 to 0.5)." We will add this more explicitly in a revised version.

8. 146, can you provide a sharper definition of "interannual extent of vegetation"? To ecologists, this may be confusing since extent almost always refers to spatial extent.

The interannual extent of vegetation refers to the variations in the area covered by green vegetation (NDVI > 0.15) over the study period (1991-2023). We apologize for the misunderstanding, as vegetation does not typically exhibit large changes in extent from year to year. What varies is whether the vegetation is greening up. To clarify the misunderstanding, we will specify and refer to it as 'interannual extent of green vegetation'.

9. 147-155. It is very difficult for readers who are not intimately familiar with the AVHRR and VIIRS datasets to follow this paragraph, even though it is quite important for the quality of the results. Terms like "flag" and "n" may be unclear. Please provide more explicit description of exactly how the monthly max/mean/min nr. of valid pixels was used and how this translates to the CI's in Fig. 2. From reading this several times I still did not understand if any correction was applied before further analysis (and looking at Fig. S1 I would expect for that to be necessary).

We plan to add a more comprehensive explanation of the procedure in the revised version:

To calculate the NDVI for each month, we started by averaging the NDVI retrievals that we obtained each month ($monthly\ NDVI = \frac{\sum_{i=1}^{n} NDVI_i}{n}$), when NDVI > 0.15. However, before 2014, the AVHRR algorithm was less strict in its data quality control compared to VIIRS from 2014 onwards, which results in more data points (n) before 2014. With n representing the total number of data points per month for NDVI calculation (see Figure S1 for n interannual variability), a higher n previously leads to lower monthly NDVI values.

To address temporal heterogeneities, we adjusted the data from the AVHRR period with the number of data points acquired during the VIIRS period. From 2014 to 2023, we identified the minimum, maximum, and average number of good quality data points for each summer month. Using these three numbers, we were able to generate a consistent variability range for calculating monthly NDVI. Hence, the NDVI values from 1991 to 2013 were recalculated by considering a similar reduction of data points as from 2014 to 2023. Figure 2 illustrates the effect of the range of NDVI values using these recalculations to estimate the interannual vegetation extent. This procedure assumes that the environmental conditions influencing the number of data points between 1991 to 2013 are similar to those between 2014 and 2023.

10. 163, it would be useful to report an accuracy metric here.
Thank you for pointing this out. We will add an accuracy metric between CARRA SWE and in situ SWE for a set of locations representing south, east and west Greenland.

11. 169, why only from January onwards and not in autumn-winter previous year?
Thank you for reflecting on the definition of rain-on-snow days. Our intention with the provided period in this indicator is to solely consider rain-on-snow days preceding the onset of the thermal growing season. This means that our aim with the chosen period is to assess whether rain-on-snow could warm the snowpack and enhance early melting.

12. 169-171, I have a slight doubt about the way that the melt rate is calculated here. If this basically represents the time that passes between the peak SWE and moment of complete snowmelt, and peak SWE occurs early in the winter-spring season, how representative is this timeframe really for the spring melt season and water release? Especially if heavy snowfall occurs later in spring and is followed by warming, this automatically leads to a situation where deep snow appear to melt more rapidly. As a reader, it is hard to fully grasp how such nuances in the choice of processing influence the results.
(Heavy) Snowfall events can indeed happen after the SWE_MAX and may slow down the snowpack melt. We acknowledge that the snowpack melting rate does not follow a linear decline. What is important to retain here is that the snowpack is providing the soil with meltwater from the moment SWE_MAX is reached, even if the snowmelt rate is not a constant as estimated. An alternative way to improve this indicator that we intend to implement for the revised version is to calculate the mean of daily mean melt rates between SWE_MAX DOY and the onset of the growing season.

13. 176, you mention rain, but rainfall does not seem to be included as a bioclimatic variable as far as I can see (Table 1, Fig. 3), while snowfall was, and rain fraction too. You refer to Fig S10 for statements on the role of rain, but this figure refers to “solid precipitation” which suggests that this is about snow. Since you discuss the role of rain regularly, why not include rain (total summer season liquid precipitation) as a bio-climatic variable explicitly? This would make your conclusions and discussion points on the role of rain more explicit and justifiable.
We described in Section 3.2 that rain is a bioclimatic indicator used to calculate rain ratio. As CARRA allows mixed precipitation, a certain volume of rain that corresponds to a low rain ratio will have a minor influence on the snowpack. Therefore, we have considered the absolute volume of accumulated rain not as relevant as rain ratio. Additionally, rain in summer is common in the southern ecoregions and less

14. You could statistically back up your choice for PCA and its assumption of linear relations. You could do this by reporting axis lengths, for instance.

Thanks for the remark! We will add to Section 3.4: As we standardized all variables prior to analyses, we opted for unimodal and linear species response, as PCA is better suited for low variance, small gradients and more intuitive for the interpretation of the biplots.

15. Fig. 1, here results are presented, and completely new information comes in (NAO / GBI), so perhaps the figure should be presented later, in the results. I also miss a scale bar for greenness and it is unclear what “greenness” represents here (is this one the extent variables you calculated, or a mean, and are pixels < 0.15 included or not?).
We apologize for the misunderstanding. Monthly NDVI in Figure 1 is described in Section 3.1. Hopefully the explanation now on point 9 facilitates the interpretation of Figure 1. We opted to remove the values of the scale from these subpanels since they correspond to a 32-year average as it would not be useful for any further interpretation. However, the figure can be simplified and split into two. The delineation of the ecoregions and the greenness evolution during summer will remain in the Methods, as they will support the readers to understand the geography of the ecoregions and to recognise the greenness dynamics across Greenland from June to August. The remaining subpanels can be moved to Supplementary Material. That way we hopefully will increase readability.

16. 227-247 seem to be combined methods and results. The source for the climate oscillation data, and the rationale for including them, have not been properly covered earlier in the methods. It is also unclear how the use of oscillations relates to your study aim and research questions.
We argue, that the paragraph mentioned is a description of the variability among bioclimatic indicators across ecoregions. We find this section in the Methods to be the most suitable place to briefly make the readers aware that there is substantial variability among bioclimatic indicators across ecoregions by indicating a few statistics, and then in the Results we actually dive into the goals of the paper.

17. 250, Pedregosa et al does not seem like the most appropriate reference for the use of PCA. I advise to find papers that specifically deal with the considerations and strengths of using PCA in a pixel-based remote sensing context.
Thank you for the comment. We will add key references there (e.g., Pearson 1901, Lorenz et al. 1956) to refer to PCA and keep Pedregosa et al. for the sake of reproducibility.

18. 259, the use of Mann-Kendall tests is state of the art, but it appears that later on you only show results for growdays and greenness, not all bioclimatic factors as suggested here? Perhaps mention only growdays and greenness then?
We performed regression and the Mann-Kendall trend test to all bioclimatic indicators in study. However, due to limited space, we only displayed GrowDays and Greenness in the main manuscript. Although other bioclimatic trends are not shown their results are referred to throughout the manuscript (e.g., Fig S2 and S10). We can specifically add the name of the bio-climatic indicators used in Section 3.4 that

supported the interpretation of the results. This will include the additional supplementary figures used to back the results as requested in point 23.

19. 262, please explain the use of a 90% confidence interval rather than 95%. With the vast amount of pixels at your disposal, and the relatively long timespan of the study, I would expect that the generally accepted 95% CI would be fine and I would be curious to know why you deviated from this standard.
Given the power of the Mann-Kendall test for a limited sample size and the associated temporal variability, we applied the test level at 10% in order to increase the probability of correctly rejecting the null hypothesis, when it is false. However, we will decrease the test level to 5%, as commonly used in ecology.

20. 271-273, the statements made here need backing; how did you test whether significant long-term trends in vegetation extent were evident? Mann-Kendall test? Could sensor discrepancies play a role here?
Thank you for your comment. We will provide the results the Mann-Kendall trend test for vegetation extent separated for the AVHRR and VIIRS periods as well as the combined period in a (supplementary) table of the revised manuscript. Although the VIIRS period is rather short, we can nevertheless apply a statistical test to assess whether the two (independent) samples are significantly different, as we did for the interannual explained variability of the principal components across ecoregions (Fig. S4).

21. 275-279, reads like methods and introduces a whole new aspect of the methodology. I would also provide some more explanation of why the use of detrended Pearson correlations is an appropriate method to evaluate linearity assumptions for a PCA.
Thank you for your comment. Pearson correlation will be appropriately introduced in Section 3.4 of the revised manuscript. An explanation of why the use of detrended Pearson correlations is an appropriate method to evaluate linearity assumptions for a PCA will be included in the revised manuscript: Pearson correlation assumes that the data are stationary; that is, their statistical properties do not change over time. Therefore, detrending supports in transforming a non-stationary time series into a stationary one, where mean, variance, and autocorrelation structure should not change over time.

22. 290 & 296, you describe how specific variables were removed from analysis a priori. This is essential information that should go into methods, and it seems at odds with your earlier statement that variables were excluded from PCA based on contribution to cumulative explained variance. I would recommend to present a single, unambiguous criterium for the inclusion of variables into PCA and figures, in the methods. Especially since the identification of useful bio-climatic indicators was an explicit aim of the study.
Thank you for your comment. We will improve our wording. Our approach was first to calculate and analyse Pearson correlations among linearly-detrended bioclimatic indicators (Section 4.1). The strong Pearson correlations are described, and the weak

linear correlations (absolute linear correlations lower than 0.3) are removed before performing PCA. In this way, we assure that the bioclimatic indicators considered meet the linear relationship criterium for PCA. Also, bioclimatic indicators with redundant information (e.g., SnowDays and GrowDays or SoilWater and SoilIce in the same season).

23. 291-292 & L. 294-295, examples of interpretation of results, and no backing (figure, reference) provided to support these interpretations.
Thanks for pointing this out. This stems from an attempt to keep the manuscript concise. While we have the results ready, we did not include all of them in the manuscript. We will make sure to include the relevant information in the Supplementary Material.

24. 311-314, I had to read this section a few times to understand the rationale and approach. So if I read correctly, you applied the PCA for all years and ecoregions separately, and then tested whether the variances explained by PC1 and PC2 were similar across the two time periods. I am not fully sure how this would demonstrate that the two NDVI records are comparable and valid in this context. The variances may be similar, but the greenness dynamics, and the associations between different variables and PC axes may not be (do I understand this correctly)? Sidenote: a lot of this information again reads like methods and not results.
Thank you for the remark. The PCA is applied to all years for the main interpretation of the results (Fig. 4). In addition, and in order to assess whether changes in the satellite sensor impact the PCA result in Figure 4, we applied PCA interannually (Fig. S4). If the interannual variability of Greenness would not be properly retrieved by both sensors, its co-variability with the remaining components would be affected and reflected in the explained variance of each component. As we do not detect significant changes in the explained variance of the first two components, we inferred that the PCA outcome is comparable despite the sensor shift.

25. This is a nice figure! Also here, a scale bar for greenness would help the reader understand what kind of magnitudes we are talking about, across regions.
Thank you for appreciating our charts! We will add a scale bar in each subplot to display the range of greenness for the years between 1991 and 2023.

26. 318-319, “PC2 is heavily shaped by continentality, permafrost extent and precipitation patterns, meaning that snow-related indicators, like SWEMAX and MeltRate have the highest explanatory power”. I struggle to see how your variables and methods could allow you to conclude anything about continentality or permafrost. This needs to be either backed up better, or (ideally) kept for the discussion. I also do not see how this means that snow related indicators are most important (snow is something different than permafrost and continentality?).
Our interpretation is based on the calculated climatology maps for vapour pressure deficit (measure of continentality), soil ice (measure of permafrost) and snowfall patterns from several seasons. The snowfall patterns are shaped by continentality as

snows amounts are relatively low (e.g., the interior of ecoregion 2). We will rephrase the above statement and move our interpretation to the discussion.

27. L 320 – 350 are altogether quite speculative and many of the claims here need to be supported either by a figure, statistics or literature (and in the latter case, it is better suited for the discussion). I would advise to back up your statements much more. And please carefully evaluate whether reported drivers are really drivers, or just represent the overall role of warming (e.g. increases in rainratio cannot really be teased apart from warming effects so I do not see how you would attribute change to rainfall patterns specifically, especially if total rainfall is not included in the analysis). I think this paragraph needs a thorough rewriting.
Thank you for your comment. We will thoroughly review our manuscript and back up our statements with evidence, either from our results or the literature. We will also rewrite the paragraph mentioned here.

28. 349-350, please consider how this relates to the aims of the study (oscillations are not introduced anywhere), report the approach in the methods, and report the test statistics either here or in the appendix.
Thank you for the comment. We will add the relevance of climate oscillations to the study in a revised version, stressing the weather pattern configuration that promoted the enhanced atmospheric warming of the past decades in Greenland. These weather pattern configurations coincide in a simplified manner with indices such as the North Atlantic Oscillation Index and the Greenland Blocking Index. In our view it was important to assess whether a certain principal component was largely shaped by climate oscillations.

29. 380, at this point the different terms used (here: spectral vegetation expansion) become a bit confusing. It would be nice to have a single, consistent term for each of the various manifestations of greening that you study in this paper, and present all of these early on.
Thank you for the notice. We will carefully and consistently review the used terms related to temporal and spatial vegetation change.

30. 382, see also major comments, here I was very unsure whether the differences in bandwidths and quality filtering might introduce artefacts into the comparison. Perhaps also good to remind the reader that 'greenness' here refers to the 0.15 threshold (related to comment above).
Thank you for the comment. We will remind the reader that Fig. 6c refers to years where the greenness threshold is met. We tested the sensitivity of the sub-periods' interval and the pattern in Fig. 6c did not significantly change, especially on showing that part of the greenness extent in ecoregion 2 did not change from the first to the second sub-period.

31. 417-420, How can you demonstrate that soil ice has an additional role, additive to warming and rainratio? Aren't they just all sides of the same coin? Could it also be, for instance, that the northern regions still feature most frozen ground conditions in

summer and that in southern regions, soils were already mostly above 0 degrees in the summer season, and that hence this dynamic is mostly evident in northern regions? I would carefully read this part of the discussion and evaluate which claims can be made with certainty, and which ones just reflect collinearity within the bio-climate variables.

The increase in SoilIce thawing in the northern ecoregions (latitudes above 70°N) is a combined consequence of warm and moist air advected from lower latitudes that contribute to warm and melt the surface (Niwano et al., 2021). Therefore, reductions in SoilIce in these ecoregions allow vegetation expansion that would not be possible with frozen soil. This is evident from the literature (e.g., Schmidt et al., 2019) that shows the rather reduced thermal growing season in 2018 with limited warm and moist air from lower latitudes.

32. Overall, the discussion would really benefit from a thematic subdivision, for instance into different sets of climate variables, or into driving mechanisms and a section on how they differ among regions? Right now the reader easily gets lost between different lines of argumentation.

Thank you for your suggestion. We will include more subsections in the Discussion to keep the line of argumentation clear and streamline the content.

33. 426 – 435, I found the descriptions of slower melt of shallower snowpacks very difficult to follow (and frankly, counterintuitive, but then I am not a snow physics expert). Even if the melt rate is lower, wouldn't the timing of complete snowmelt still be earlier for shallow snow than for deeper snow? What then is the exact role of the slower melt rate and potentially better water absorption within the context of your findings? I have a feeling that similar claims could be made about the role of deeper snow and its impact on soil temperature and microbial activity (as you also state in the introduction), so I am still in the dark about the role of melt rate in nutrient availability. I would recommend rewriting this in a way that is more accessible to readers without a background in snow physics and staying closer to your own results.

We hope that the improved explanation on point 4 clarifies the relationship between snow depth and snowmelt rate better. The referee is correct: in addition to slower snowmelt, the onset of the thermal growing season is generally occurring earlier. Instead of a quick snowmelt water runoff from relatively deeper snowpacks, the slow snowmelt water percolation in the soil can be used not only for microbial activity to generate nutrients but also directly by vegetation.

34. 428, Heijmans et al (2022) doesn't deal with the release of nutrients in relation to spring water availability. Perhaps we cite others in our review that have relevant findings on this topic, but to me this doesn't seem to be an appropriate reference here.

Thank you for the remark. Heijmans et al (2022) is a very relevant reference for us to better understand the links between tundra vegetation and permafrost, but indeed we should have cited the work of Salmon et al. (2016).

35. 463-465, maybe you can back up this hypothesis about the role of shrubs or potentially other species groups by checking your greenness trends against the CAVM or Karami et al. (2018)?
The recommended sources, the Circumpolar Arctic Vegetation Maps and Karami et al. (2018), are static maps based on the collection of data over several years with different approaches. We see the potential of the recommendation, but it is enough work for a publication on its own.

36. 475, what exactly do you mean by "validating bio-climatic indicators"? I think you could explain your proposed course of action a bit better, and also explain how that would help understand future trends.
Our study demonstrates how the co-variability among bio-climate indicators effectively clusters greenness. It is crucial that these indicators are accurately modelled in global climate models. If the historical period in global climate models does not accurately reflect the conditions of our study period, these models will be ineffective at predicting future spatial and temporal vegetation changes in the ice-free regions of Greenland. We will revise the manuscript to make this clearer.

37. The implications section reads like a rather surprising selection of several implications, of which I am not really sure if all the main ones are represented, and whether the ones that are now discussed most extensively are in fact the most important ones. For example, a lot of attention is dedicated to PBAPs and fog, but no mention is made of carbon dynamics or surface energy balance feedbacks. Even if this is deliberate, it would be good to highlight why specific implications are discussed while others are not. You do mention some of these aspects in the limitations, but they are of course also relevant from an implications perspective.
Thank you for the remark! We find relevant to keep recent literature that links primary biological aerosol particles (PBAPs) with the cloud formation in the Arctic and the potential of generating fog conditions due to decreasing sea ice as part of the Discussion. We mentioned a few other implications although not directly such as the feedback of the vegetation canopy on the surface feedback, shifts in cloudiness due to increased PBAPs, the cooling of the surface due to surface evaporation and ecological shifts on the animal community. However, we acknowledge that other important implications such carbon dynamics and surface albedo feedback have not been addressed in detail and we will include those in the revised manuscript.

38. 506 – 510, I would expect that such episodes of warm, humid conditions should be evident from your PCA analysis, so I do not see the point of mentioning the role of this particular episode as a limitation?
We agree that this was mentioned in a confusing manner, we will rewrite accordingly.

39. 517-520, needs references for the claims made. I would like to add that while permafrost thaw can indeed release moisture or lead to ponding, deeper thaw fronts also often lead to deeper infiltration and surface drying (Liljedahl et al., 2016). This section could use more nuance and backing.

Thank you for mentioning the possibility of deep thaw fronts which could lead to deep infiltration and surface drying as described by Liljedahl et al. (2016). Indeed, the complexity of these feedbacks is high and we will provide a more nuanced perspective in a revised version.

40. 525 – 530, I do not want to send you back to the drawing board, but I am interested why elevation was added to your analysis, while aspect and slope were not. You rightfully stress their importance and I would (perhaps naively!) assume that it would not be such an enormous effort to include them in your analysis as well?
Good point and indeed, for a wide perspective it would be useful to add slope and aspect for which we have the data and the analysis ready. We will add these results to the revised version.

41. Rather than reiterate what you did, you could summarize the actual findings and try to align better with the original aims (perhaps mention which set of variables or which variables show the strongest associations?) and mention the key advance you have made? This would make the conclusion more informative.
Thank you for the suggestion. We will briefly summarize our findings and keys advancements in the field in the revised version.

Technicalities & Language

1. 10 “summer spectral vegetation”. This is an unusual term, it would be good to rephrase it or explain it so that there can be no ambiguity about what it means.
2. 18 “by 22.5% increase” should be “by 22.5%”. I also recommend to be more explicit about what you mean by “the distribution of vegetation”. Do you mean that the vegetated area of Greenland (determined here as summer NDVI > 0.15?) expanded in area by 22.5%? Perhaps you want to rewrite this sentence.
3. 25, what do you mean by “regional Greenland”? Perhaps that specific regions of Greenland are warming three times faster.
4. 31, add “and” instead of comma between “composition” and “alterations”.
5. 48, is it really necessary to mention the specific methods of Gamm et al (”using [..], [...] and [...]”)? This is not done for other papers that you cite?
6. 53-55, this reads like a repetition of L. 43-44.
7. 62, I do not think “snow cover melt” is a very generally used term. Maybe write “snowmelt timing” or “snow melt rate”, depending on what you mean exactly?
8. 71, maybe write “large amounts of snow” rather than “large amounts of snow coverage”, since from what I understand snowpacks were also very deep, not just spatially extensive.
9. 81-82, example of a grammatically confusing sentence.
10. 83-86, implications for phytoplankton seem beyond the scope of your study system and I do not see the added value of discussing it here (it seems more of an implication rather than an example of the importance of subsurface flow to terrestrial vegetation).
11. 105, add “the” between “to” and “CARRA”.
12. 132, “and thereafter is then continued” should be “and is thereafter continued”.
13. 133, add “is” between “mask” and “spectrally”.

14. Figure 5) Final sentence in the caption: Do you mean that the trend was considered significant if the 90% CI of the estimate did not overlap 0? This is what I am used to. Similar for Fig. 6
15. 376, replace “evidence” with “shows”?
16. Table 2) perhaps a no brainer, but it would be good to explain what the fraction mean; is this % of total area of that ecoregion?
17. 446, change “favourable areas” into “a more favourable area”.
18. 498, change “as” into “as in”

Thank you for these valuable edits which we will incorporate in the revised manuscript.

References

Martin, A. C., Jeffers, E. S., Petrokofsky, G., Myers-Smith, I., & Macias-Fauria, M. (2017). Shrub growth and expansion in the Arctic tundra: an assessment of controlling factors using an evidence-based approach. Environmental Research Letters, 12(8), 085007.

Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Normand, S., Rüger, N., Beck, P. S., ... & Weiher, E. (2018). Plant functional trait change across a warming tundra biome. Nature, 562(7725), 57-62.

Miura, T., Turner, J. P., & Huete, A. R. (2012). Spectral compatibility of the NDVI across VIIRS, MODIS, and AVHRR: An analysis of atmospheric effects using EO-1 Hyperion. IEEE Transactions on Geoscience and Remote Sensing, 51(3), 1349-1359.

Liljedahl, A. K., Boike, J., Daanen, R. P., Fedorov, A. N., Frost, G. V., Grosse, G., ... & Zona, D. (2016). Pan-Arctic ice-wedge degradation in warming permafrost and its influence on tundra hydrology. Nature Geoscience, 9(4), 312-318.

Mekonnen, Z. A., Riley, W. J., Berner, L. T., Bouskill, N. J., Torn, M. S., Iwahana, G., ... & Grant, R. F. (2021). Arctic tundra shrubification: a review of mechanisms and impacts on ecosystem carbon balance. Environmental Research Letters, 16(5), 053001.

References:

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy,W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K. P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D. S., Subias, A., Tijm, S., Toll, V., Yang, 575 X., and Køltzow, M. Ø.: The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system, Monthly Weather Review, 145, 1919–1935, https://doi.org/10.1175/MWR-D-16-0417.1, 2017.

Liu, Y.,Wang, P., Elberling, B., and Westergaard-Nielsen, A.: Drivers of contemporary and future changes in Arctic seasonal transition dates for a tundra site in coastal Greenland, Global Change Biology, 30, e17 118, https://doi.org/10.1111/gcb.17118, 2024

Lorenz, E. N. (1956). Empirical orthogonal functions and statistical weather prediction (Vol. 1, p. 52). Cambridge: Massachusetts Institute of Technology, Department of Meteorology.

Madson, A., Dimson, M., Fortini, L. B., Kawelo, K., Ticktin, T., Keir, M., ... & Gillespie, T. W. (2023). A near four-decade time series shows the Hawaiian Islands have been browning since the 1980s. Environmental Management, 71(5), 965-980. https://doi.org/10.1007/s00267-022-01749-x

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., ... and Voldoire, A. (2013). The SURFEXv7. 2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. Geoscientific Model Development, 6(4), 929-960.

Musselman, K. N., Clark, M. P., Liu, C., Ikeda, K., and Rasmussen, R.: Slower snowmelt in a warmer world, Nature Climate Change, 7, 214–219, https://doi.org/10.1038/nclimate3225, 2017.

Myers-Smith, I. H., Kerby, J. T., Phoenix, G. K., Bjerke, J. W., Epstein, H. E., Assmann, J. J., John, C., Andreu-Hayles, L., Angers-Blondin, S., Beck, P. S., Berner, L. T., Bhatt, U. S., Bjorkman, A. D., Blok, D., Bryn, A., Christiansen, C. T., Cornelissen, J. H. C., Cunliffe, A. M., Elmendorf, S. C., Forbes, B. C., Goetz, S. J., Hollister, R. D., de Jong, R., Loranty, M. M., Macias-Fauria, M., Maseyk, K., Normand, S., Olofsson, J., Parker, T. C., Parmentier, F.-J. W., Post, E., Schaepman-Strub, G., Stordal, F., Sullivan, P. F., Thomas, H. J. D., Tømmervik, H., Treharne, R., Tweedie, C. E., Walker, D. A., Wilmking, M., and Wipf, S.: Complexity revealed in the greening of the Arctic, Nature Climate Change, 10, 106–117, https://doi.org/10.1038/s41558-019-0688-1, 2020

Niwano, M., Box, J. E., Wehrlé, A., Vandecrux, B., Colgan, W. T., & Cappelen, J. (2021). Rainfall on the Greenland ice sheet: Present-day climatology from a high-resolution non-hydrostatic polar regional climate model. Geophysical Research Letters, 48(15), e2021GL092942, https://doi.org/10.1029/2021GL092942

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin philosophical magazine and journal of science, 2(11), 559-572 , https://doi.org/10.1080/14786440109462720

Pourmohamad, Y., Abatzoglou, J. T., Belval, E. J., Fleishman, E., Short, K., Reeves, M. C., ... & Sadegh, M. (2024). Physical, social, and biological attributes for improved understanding and prediction of wildfires: FPA FOD-Attributes dataset. Earth System Science Data, 16(6), 3045-3060. https://doi.org/10.5194/essd-16-3045-2024

Salmon, V. G., Soucy, P., Mauritz, M., Celis, G., Natali, S. M., Mack, M. C., & Schuur, E. A. (2016). Nitrogen availability increases in a tundra ecosystem during five years of experimental permafrost thaw. Global Change Biology, 22(5), 1927-1941.

Silva, T., Abermann, J., Noël, B., Shahi, S., van de Berg, W. J., and Schöner, W.: The impact of climate oscillations on the surface energy budget over the Greenland Ice Sheet in a changing climate, The Cryosphere, 16, 3375–3391, https://doi.org/10.5194/tc-16-3375-2022, 2022.

Schmidt, N. M., Reneerkens, J., Christensen, J. H., Olesen, M., and Roslin, T.: An ecosystem-wide reproductive failure with more snow in the Arctic, PLoS Biology, 17, e3000 392, https://doi.org/10.1371/journal.pbio.3000392, 2019.

Schyberg, H., Yang, X., Køltzow, M., Amstrup, B., Bakketun, m., Bazile, E., Bojarova, J., Box, J. E., Dahlgren, P., Hagelin, S., Homleid, M., Horányi, A., Høyer, J., Johansson, m., Killie, 750 M., Körnich, H., Le Moigne, P., Lindskog, M., Manninen, T., Nielsen, E. P., Nielsen, K., Olsson, E., Palmason, B., Peralta, A. C., Randriamampianina, R., Samuelsson, P., Stappers, R., Støylen, E., Thorsteinsson, S., Valkonen, T., and Wang, Z.: Arctic regional reanalysis on single levels from 1991 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.713858f6, accessed on 15-12-2022, 2020.

Skakun, S., Justice, C. O., Vermote, E., and Roger, J.-C.: Transitioning from MODIS to VIIRS: an analysis of inter-consistency of NDVI data sets for agricultural monitoring, International Journal of Remote Sensing, 39, 971–992, https://doi.org/10.1080/01431161.2017.1395970, 2018.

Vermote, E., Justice, C., Csiszar, I., Eidenshink, J., Myneni, R., Baret, F., Masuoka, E., Wolfe, R., Claverie, M., and Program, N. C.: NOAA Climate Data Record (CDR) of Normalized Difference Vegetation Index (NDVI), Version 5, https://doi.org/10.7289/V5ZG6QH9, access date: 2022-05-06, 2018.

Vermote, E., Franch, B., Roger, J.-C., Murphy, E., Becker-Reshef, I., Justice, C., Claverie, M., Nagol, J., Csiszar, I., Meyer, D., Baret, F., Masuoka, E., Wolfe, R., Devadiga, S., Villaescusa, J., and Program, N. C.: NOAA Climate Data Record (CDR) of Surface Reflectance, Version 1, https://doi.org/10.25921/gakh-st76, access date: 2023-07-06, 2022.

Wilks, D. S. (2011). Statistical methods in the atmospheric sciences. Academic press.