Comments to the author:
Dear Dr. Löffler and co-authors,

I have now received the report from an additional reviewer, given that the previous reviewers indicated that they did not want to re-assess your manuscript. The reviewer points out that your paper needs some further improvement and makes some excellent suggestions, which should help improve the clarity of your paper. Please
1) explain more clearly how the statistical analyses relate to the specific questions asked,
2) structure the discussion around the main questions posed in the introduction,
3) outline more prominently the limitations of the study,
4) use consistent terminology throughout the manuscript and
5) take into account the additional specific comments of the reviewer.

I look forward to seeing your revisions. Given that the review process of your paper has been exceptionally lengthy I point out that the revised manuscript will be reconsidered only if your revisions address the issues raised in a profound and conclusive manner.

Best regards,
Michael Bahn

Many thanks to the reviewer and the editor for their detailed comments and suggestions, which helped a lot to further improve the manuscript.
In accordance with the suggestions, we again revised the manuscript substantially. Specifically, we 1) improved on some methods used and rewrote the text on the statistical analyses in the Material and Methods section. Furthermore, we 2) restructured the Discussion around the main questions posed in the introduction and 3) included a paragraph on the limitations of the study. We additionally revised 4) the terminology throughout the manuscript, and 5) addressed the detailed comments by the reviewer.

Review for Dobbert et al. The application of dendrometers to alpine dwarf shrubs – a case study to investigate stem growth responses to environmental conditions

The manuscript of Dobbert et al. deals with intra-annual growth dynamics of alpine dwarf shrubs obtained by dendrometer measurements. The applied methodology is new since dendrometer measurements on shrub species have not been conducted so far. The topic is important since growth dynamics at the alpine treeline are changing due to climate warming. The manuscript addresses the eco-physiological mechanisms of growth which might help to understand dynamics that are observed at the alpine treeline.

The manuscript has received already a lot of comments through two rounds of major revisions, which have been addressed in the revised version and should be

acknowledged. However, I still feel that improvements could be made, particularly in relation to the methodological description and discussion.

Methods: I have spent some time trying to understand what kind of statistical analyses have been made, and why. However, I could not fully reconstruct what was done in detail, since analyses are made with a lot of different response variables and explaining factors, and on different time scales (annual values, daily values). For example, I did not understand why a multiple regression analysis was done and afterwards a prediction from a linear regression model. Further, why didn't you use a mixed effect model with site as random factor since all sites have been measured four times. It would be helpful to clearer describe the statistical methods and what kind of research question should be answered with the respective analysis. See also detailed comments.

We agree that the Methods section (especially chapter *2.5 Climate-growth relations and potential drivers of radial stem change,* on the statistical analysis) needed some clarifying. As suggested, we now included linear mixed effect models with site as random effect (see also our response to the detailed comments below). In general, our statistical analysis followed the following structure:

1. Identification of growth defining parameters in order to assess major growth patterns and their variation between years and specimens. We here used Gompertz models to accurately define critical dates of growth. Subsequently, we used linear mixed effect models (in the revised version) to assess the relation of these growth-defining parameters and their importance for total annual growth.
2. Linking the previously defined annual parameters to environmental parameters derived from raw environmental measurements. This was done using a simple correlation analysis, with the aim to identify the most important environmental drivers controlling growth patterns.
3. Examining intra-annual patterns of stem diameter shrinking and swelling in response to fine-scale, local environmental conditions (daily values). Here, we performed moving window correlations.

To convey this governing structure more clearly to the reader, we substantially revised the Material and Methods chapter. In chapter *2.5 Climate-growth relations and potential drivers of radial stem change*, we restructured the text according to the three main methods deployed and made sure to point out the aim of the respective analysis in relation to the research questions introduced in the Introduction.

Discussion: You define three research questions at the end of the Introduction. I think it would help the reader to build-up the discussion along these three main questions, or at least summarize in the beginning of the discussion the main findings of the study, which will be discussed then. At the moment I have the feeling that the discussion is not well structured and is not based on the findings that were presented in the results section. For example, you start the discussion with a topic that was even not shown in the results (zero-growth years), and you refer to new Figures in the Supplement. I would recommend to discuss the findings that were presented in the results section.

Also, I have the impression that patterns of stem growth and stem diameter changes (including stem water relations) are sometimes mixed-up, especially within the discussion.

In accordance with your detailed suggestions we restructured the Discussion section entirely. After a short paragraph focusing on the methodical implication of the study we now present three subchapters, each focusing on one of the questions posed within the Introduction. Also, we moved all parts presenting additional results from the supplementary material to the Results section and made sure that all findings prominently discussed in the Discussion are also clearly mentioned in the Results. Additionally, we paid special attention to our use of the terms stem growth and stem diameter changes (as well as similar terms).

Further, I think the limitations of the study as nicely summarised by the previous reviewer, should be placed more prominent within the discussion.

After restructuring the Discussion chapter, we now start with a clearer focus on the methodical implications of our study and included the limitations in a separate paragraph at the beginning of the Discussion. Here, we discuss the limitations mentioned by the previous reviewer, including the challenge of extracting physiologically meaningful data, potential intra-plant variation, and the potential of including VPD as a critical environmental factor.

Finally, terms should be used consistently, e.g. you use various different terms for stem diameter variations (stem diameter change, stem diameter data, stem change, stem diameter variability /variation, annual stem variability)

As suggested, we refrained from using multiple terms here and consistently use "stem diameter variability" to refer to the measured stem diameter variations. However, when we refer to the total change in stem diameter from the start to the end of the year (stem diameter at the end of the year - stem diameter at the beginning of the year) we consistently use the term "stem diameter change" to differentiate the two.

Detailed comments:

L16: can you explain what growth-defining parameters are? Do you refer to phenology?

We here referred to the parameters used throughout the study to define radial stem growth. They are defined in chapter *2.3 Analysis of seasonal growth patterns,* in the Material and Methods section. As this might not be clear to the reader while reading the Abstract, we use the term "radial stem growth" in the revised version.

L16: response of what?

To clarify, we changed this sentence. It now reads:

> "We found high inter-plant variability in overall radial stem growth, but strong similarities in response patterns to the local environment.".

L18: remain

We believe "remains" is the grammatically correct term here, as we refer to a singular evergreen species (*E. hermaphroditum*).

L18: Why can we conclude this from a dependency of growth start to winter temperature and ground freezing?

We reworded this part of the Abstract as it might not be clear in the previous form. It now reads as follows:

> "Our results suggest that the evergreen species is highly adapted to the specific local conditions, remaining partly photosynthetically active during the snow-free winter, which facilitates carbohydrate accumulation for early-season physiological activities."

The conclusion that *E. hermaphroditum* might stay partly photosynthetically active during the winter months stems from a combination of several aspects of our results: Due to the snow free conditions at the monitored ridge positions, the evergreen species is faced with unique environmental conditions, receiving continued solar radiation which might force photosynthetic activity. At the same time, we found the species capable of avoiding negative effects of extreme temperatures through a process of radial stem shrinking, and an early growth start was linked to low winter temperatures, suggesting *E. hermaphroditum* is adapted to profit from cold winters with associated high solar radiation. This is discussed in detail in the Discussion section.

L23: Can you be more precise? What does this finding mean for the debate on greening and browning?

As detailed in the final paragraph of the Discussion, our findings regarding drivers of shrub growth have important implications for future distribution and potential spread/decline of the species, which in turn influences the processes of arctic greening and browning. We therefore included this in the Abstract:

> "We identified soil moisture availability and winter freezing conditions as the main drivers of stem diameter variability, which might negatively affect the species distribution in a warming climate, thus forwarding the ongoing debate on the functional mechanisms and complexity of vegetational shifts in arctic and alpine regions. "

This also ties back to the first sentence of the Abstract.

Methods:

L138: I don't see this in Fig. A3

Fig. A3 (a) illustrates the dendrometer set up as described here, showing that the dendrometer was not installed inside the radius of other larger shrub species or near large stones. Because of the complex set up and vegetation in the field it was challenging to show a good photo here.

L143: can you explain briefly the daily mean approach? Why didn't you use the daily maximum?

Originally, we tested both approaches for our data, calculating daily mean and maximum values (calculated from 24 hourly measurements for each day, from 0:00 to 23:00). For the purpose presented here (annual growth patterns), we found the results highly similar, which is in accordance with results by Deslauriers et al. (2007). They found differences in both approaches mainly in daily amplitude, which was not of interest for our study. Consequently, we chose the daily mean approach, believing it to be the simpler of the two. Yet, considering the zero growth approach deployed later in our study for calculating total annual growth, the daily maximum approach might be slightly more appropriate here, which is why we changed this in the revised manuscript and added a short explanatory sentence in the text. As evident from Figure 1, this change in methodology resulted in only minor changes in the resulting curves.

L173: I do not understand why you derived the dates from the modelled curve and not from your original (growth) data?

Using sigmoid models to determine growth start and cessation has proven successful in a number of past studies, especially, since these models can be biologically interpreted (e.g. Duchesne et al., 2012; Van der Maaten et al., 2018; Liu et al., 2019). Due to the additional measuring of reversible stem hydraulic dynamics by dendrometers, raw dendrometer data are known to deliver rough estimates of cambial activity and less reliable critical dates like onset and cessation of growth (Deslauriers et al., 2007; van der Maaten et al., 2018; Cruz-García et al., 2019). This is especially true for our measurements of shrub growth, because, in comparison to trees, they showed a less clearly defined growing season, with a phase of stem shrinking during the winter months, and consequent stem expansion in spring. The chosen modelling approach therefore proved the most reliable to clearly define the main growing season, derive the critical dates and ensure high comparability between years and specimens. We included this statement in the text.

L174: Why did you chose a threshold of 20% as growth start, this is quite a lot. How was the 'careful testing' done and which criteria did you use to validate your threshold. In Fig. 1 it looks like that the growth season starts when the growth rate is already close to its maximum.

Our testing was mainly done visually by direct comparison of raw environmental data (for each specimen) and the resulting growing season. One of our main concerns here was to make sure that the defined growing season does not include radial stem expansion resulting from the stem shrinking observed during the winter months and probably linked to stem water dynamics, which is why we initially chose a higher threshold for growth start. However, since the difference between a 10 % and a 20 % threshold was minor, we changed it in the revised version. In Figure 1, we included lines to aid the visual interpretation of the figure.

L216: can you give the response variable (annual growth and peak shrinking according to Table 1) and the independent variables? In table 1 only 3 independent variables are given, in Line 219 you write six measures, please clarify.

In Line 219 we refer to the six measures of collinearity, which are implemented in the mctest package (Imdad and Aslam, 2018), and which we deployed to assess collinearity between independent variables. We revised this entire chapter and performed linear mixed effect models, replacing the linear regression (see below).

L223: Is it correct that you have N=12*4=48 for the correlations? Why didn't you use mixed models with sites as random effects? How did you account for the site effect in your analysis?

This is correct. We agree that mixed models might be a better fit here and replaced this analysis in the revised version. We now used linear mixed effect models with site (including study region and elevation) as random effects. The results are presented in Table 1.

L223-25: I do not understand what was done here. What was the research question for this analysis and what is the difference to the correlation analysis that was done before?

The results of this analysis (formerly presented in Fig. 4) is indeed to some extent redundant with the results from the correlation analysis. We therefore excluded it in the revised version.

L224: What window size did you chose and did you test for the effect of the chosen size?

We did indeed test several values for window widths (presented in an earlier version of this manuscript) and finally chose 3 days, which showed the overall most significant results. We included this in the revised text.

Results:
L238: Since you do not describe solely growth pattern, I would use stem diameter variation instead

We chose to use "stem diameter variability" here (see above).

L244: What do you mean by timing of shrinking period? Is it the start or the peak or end?

The start of the shrinking period. We included this in the text.

L259: What do you mean with total growth pattern? Did you test these statements statistically? I do not think that year-to-year variability is similar between the two regions. In region 1, the year 2017 has lowest growth, while the other three years are comparable, in region 2, growth is exceptional in 2016, while it is low in the three other years.

We do agree that this statement should be more specific. We changed it to "...growth patterns and timing of the growing season were similar in the two studied regions", which is more accurate and shown in Figure A4 (e.g. late growth start in 2015). Here it is also important to keep in mind that Fig. A4 shows averaged curves while variation between specimens (indicated by standard deviation in the figure) in general higher than variation between regions, as indicated in the text.

L261: but here (Fig. A1) you do not look at growth patterns, but at stem diameter variation.

We do not refer to Figure A1 in this line. Instead we here refer to Figure A4. We agree, however, that we do look at stem diameter variation, rather than growth. We therefore changed the term here to " patterns of stem diameter variability".

Fig. A4 and A6: It looks like
West and East is swapped (A4: generally higher annual growth for West, and high increment in 2016 for East, in A6 the other way around).
Thank you very much for pointing that out. We revised Fig. A4 accordingly.

Table 3: what do you mean by standard deviation , +/- SE: SE is the abbreviation for standard error, what was calculated? I do not think that these values are meaningful, because they are dependent on the mean. Better would be to calculate e.g., coefficient of variation, if you are interested in the variability of the measure.
Originally, we included the standard error (SE) here, as a measure of variation. However, we do agree that the coefficient of variation (CV) is the better measure here. We therefore calculated it in the revised version to show variability between sampled specimens. CV is now included for Table 1,3, and 4.

L306-308: How was this analysed and where are the results presented? Isn't the model overfitted with 25 explaining variables and only 48 observations?
We agree. The model was initially created to check the overall reliability of our chosen parameters and is complementing other results. As it is not prominently discussed in the discussion section we removed this part from the text.

L310: The moving window analysis does not present climate-growth relations, since you use stem diameter changes and not only growth. This measure has a totally different dynamic as also stem water relations are included.
We agree that climate-growth relations might be the wrong term here. We changed it to "the relation of stem diameter variability and environmental conditions", which is more accurate and in consistence with the terms used throughout the text.

Fig. 4 and 6: What is the difference of the results presented here? Why is annual growth related to radiation and soil moisture in Fig.4, and to soil moisture maximum in Fig. 6, and not at all to radiation?
The analysis presented in Fig. 4 does include different parameters from the correlations presented in Fig. 6. However, the multiple regression analysis is not very robust compared to the correlation coefficients, and the results to some extent redundant. We therefore removed Fig. 4. We revised the text (Methods, Results, and Discussion) accordingly.

L290-300: In my opinion this belongs to discussion

We moved some of the aspects presented in this paragraph to the Discussion section.

Discussion:
L351: Do you refer here to the soil moisture shown in Fig. A1b? But these measures vary between sites and regions. In Fig. 2 it is not visible whether there are changes between sites /regions.
Here we referred to the environmental conditions characterising the exposed ridge positions, described in the previous sentences, rather than variation between sites/regions. For example, all of our study sites in both study regions and along the elevational gradient are experiencing discontinuous snow cover and high solar radiation, compared to their immediate surroundings. As we revised the Discussion, we placed less emphasis on this aspect and also made sure to include it more prominently in chapter *2.1 Study sites.*

L353: But the growth curves of the two regions look totally different at least for the years 2015 and 2018!
We here referred to seasonal growth patterns (including timing of the growing season and shrinking period in winter). However, since this is not completely clear from the graphs, we removed this part from the Discussion.

L354: but which was not tested in your study as you did not include regional climate data. If you argue like that it would be good to include the chain of arguments developed in line 290-300 here.
Since this aspect seemed to be not entirely clear from the presented results, we decided to remove it in the revised version. We did, however, include some of the aspects from line 290-300 later in the text.

L361-369: If discussed so prominently, these data should be presented in the results section.
We included these results in the Results section and discussed them less prominently in the restructured version of the Discussion.

L362: how did you link the zero growth to snow coverage periods and where can this analysis be seen?
This is presented in Fig. A5 (formerly Fig. A9). We moved this to the Results section.

L466: With which analysis can this statement be confirmed?
These results are from the correlation analysis. We revised this part, directly linking it to the results from the year 2016 (high number of days with temperatures > 5°C resulting in high growth). Also, the positive effects of winter temperatures shown in the correlation analysis are discussed earlier in the text.