

Authors' response

To bg-2021-260-RC1 (05 Nov 2021): We thank the anonymous referee #1 for their comments. We will address all **general comments** raised regarding the structure and readability of the manuscript in our next revision or resubmission. We will refine the language to reduce ambiguity and supplement our statements with appropriate citations where indicated by the referee.

As the manuscript will, hence, undergo major structural changes, we can not address all **specific / technical comments** in detail at this point. We will take all specific and technical comments which are still relevant into consideration.

In the following, we shall give a brief respond to all relevant issues.

General comments

- "the [...] study is highly speculative and not based on any experimental evidence [...]" We are aware that our work is not complemented by experimental evidence and did not claim otherwise. We welcome relevant observational data to further improve parameterizations of Arctic and subarctic species. We ought to stress that we do not own the "true" parameterization for subarctic vegetation but assess the uncertainty arising from the use of generic parameterizations.
- "Many claims sound superficial [...], such as [...] the alleged deviation of the years 2018 and 2019 from the site's climatology [...]" We strongly disagree with this statement. We assessed the particular years 2018 and 2019 by statistical means with respect to our derived climatologies. We acknowledge that we have not used common tools such as the infamous *p-value* (for a discussion see statement of the American Statistical Association (2016)).
- "The manuscript is too long and the organization is confused [...] no clear division between methodological section and results section" We will address this in the next revision.

Specific comments

- Lines 51-52. "and leads to a build-up of ozone and its precursors during winter." Are you sure? How can BVOC accumulate in the atmosphere if the vegetation is covered by snow? Please add some citations to support this claim. Much more credible is the following explanation based on stratospheric intrusions. We agree with the referee that a vegetation covered by snow would not release BVOCs to the atmosphere. We implicitly assume this. Therefore, we did not explicitly state that BVOCs are not among the ozone precursors that accumulate in the course of winter. If this connection is not obvious to the readers, we will rephrase the sentence accordingly: "[...] ozone and its non-biogenic precursors [...]"
- Line 63. "the time in which vegetation can accumulate ozone." It sounds very bad written this way: vegetation does not accumulate ozone because ozone it is not

bioaccumulative. Did you mean the dose? Thank you for pointing this out. We will rephrase accordingly and refer instead the accumulated ozone dose.

- Line 77. " $\langle [O_3] \rangle = 36\text{-}54\text{ppb}$ " Please explain the formalism. What do different brackets mean? The bracket notation commonly refers to the concentration of a chemical species. Parts-per-billion (ppb) is a volume mixing ratio (VMR) and strictly speaking no concentration. Though, they are often used synonymously because measured VMR is reported in concentrations by simply multiplying with a factor 2. So, perhaps, we should use another terminology instead, e.g. χ_{O_3} (e-education.psu.edu).
- Line 81. "A substantial body of evidence exists that suggests flux-based metrics, that relate stomatal ozone uptake to vegetation damage, are biologically more relevant for risk assessments than exposure-based metrics." Well, please cite at least some works of this "substantial body" Thank you, we will include relevant citations.
- Figure 1. This figure was never referred in the text. This Figure has been referred to in Sect. 2 L177/179.
- Figure 2. It does not seem to me that the O_3 concentrations of 2019 are different from those of 2018. The spring peak could even be identical (although unknown, because in 2019 O_3 measurements started about 20 days after the spring peak) It is remarkable how the referee can infer from Fig. 2 that ozone concentrations do not differ between 2018 and 2019. Based on the observed visible damage on ozone sensitive vegetation in 2018 but not in 2019, our working hypothesis was that either ozone or one of the other important environmental variables may have differed. In Sect. 2, we therefore show whether and how they differ. We found that there are but only small differences.
Regarding the spring peak, until the beginning of May most of the vegetation in the subarctic is still either covered by snow (see Table 3 for the dates) or bud burst has not yet occurred. Hence, it was unlikely for non-coniferous trees to experience the ozone spring peak in 2018/19. We will come back to coniferous trees later.
- Line 146. Please make clear the acronym PFT on first use Plant functional type (PFT) was introduced in the abstract. Following the common code of conduct, we ought to introduce it also in the main matter. Thank you for pointing this out.
- Line 165. "luftkvalitet.no" What is it? And EBAS? Please make them clear. "luftkvalitet.no" is the webinterface to the database of air quality measurements in Norway operated by NILU. EBAS is a database operated by NILU as well. It comprises global air quality measurements, e.g. trace gases, pollutants, particular matter from different networks. We will elaborate on this.
- Line 182. "This indicates that the vegetation was more affected by ozone in 2018 than in 2019." Being affected by visible symptoms does not necessarily mean having

suffered biomass or productivity reduction. It is correct that biomass reduction and visible damage due to ozone are not the same. We discuss this briefly in Sec. 5 but will elaborate on it to make the difference clearer.

- Line 190. "high ozone concentrations ($[O_3] > 40\text{ppb}$)" It is strange to read that O_3 concentrations above 40 ppb are "high" concentrations. 40 ppb is not particularly high. It has been the threshold for the toxic accumulated ozone exposure (AOT) on ecosystems since its introduction at the UN ECE-Workshop in Egham in 1992.
- Line 194. "A method for gapfilling data has been presented in Falk et al. (2021)." Ok, but was it then applied to this work? Please write it. We will rephrase: "A method for gapfilling data has been presented in Falk et al. (2021) and applied in this work."
- Line 205. "We evaluate the statistical significance of divergences from the norm in these variables (referred to as anomalies) in 2018/19" I suspect a misuse of the locution "statistical significance". How was this significance assessed? Which statistical test was applied? What is the level of significance? We will elaborate on the description the devised hypothesis test. First, we compute the excess number of days with residuals larger than 1σ . This means, we calculate the daily mean and std for 21 years \times 24 hours (264 data points). From this, we find the residuals in 2018/19 and divide by std. In the next step, we explicitly assume that these are normal distributed - we will elaborate on this assumption in the next revision and check the underlying distributions (but see also Fig. 4). Due to the variability in the data, we choose a low sigma level for the significance test (1σ). By grouping the data into months, we find the number of days with observations in each month. We count the number of days with residuals exceeding 1σ and compute the ratio with the number of days of observation. If the data was normal distributed, we would expect 18% of data in this interval. If the percentage of days exceeding 1σ is larger, we state that we find a significant excess of days with high ozone.
- Lines 206-208. I do not understand. Please, explicit the methodology. We will rephrase and include a brief description of the method. We use a Reynolds decomposition on the 1986-1996 data from Svanvik and data from all other long-term monitoring stations in Fennoscandia (1992-2012) to establish multi-annual means. We then remove the annual cycle from the July 2018 observations at Pallas (reference station) to find the anomalies. We have shown that these correlate best with observations at Svanvik. We find the time lag between the observations and shift the time series of anomalies accordingly. We then add the climatology of Svanvik to retrieve a reconstruction of the missing data. As 2019 does not deviated much from the norm by means of temperature, precipitation, and irradiance, we assume it for a normal year. From this, we find a probable offset of the historical ozone climatology compared to present day and correct for this. We compared this reconstruction with high resolution air quality model reanalysis and found a similar



accuracy (residuals compared to observation). For details, please refer the cited work Falk et al (2021).

- Line 213. "Averaged monthly accumulated precipitation (blue bars) is shown with standard deviation" It is not consistent to show SE once and STDEV the other time. The use of SE is more appropriate when estimating averages. We will change the figure accordingly.
- Line 228. "Darker colors indicate higher probability to observe these values."
- Line 229. "On top of the density distributions, a 10 days average of daily mean ($\overline{h[O_3]_{10d}}$) is displayed together with 1σ uncertainties and SE, respectively" What does it mean? It is not clear to me. Why show a probability density if you are plotting a multiannual average? Or does the line represent the median instead? In Fig. 4, we show the distribution of ozone data in 2-dimensions. The dots represent the daily averages (but only every 10th day is shown). We will rephrase the caption to correct for this. We show the daily average to infer that not all daily mean ozone are strictly normal distributed (tails in one or the other direction) but it is also visually deductible that most data fall in the 1σ band justifying the chosen confidence intervals for a normal distribution. We shall, however, elaborate on this by statistical means.
- Line 232. "The decline in $\overline{h[O_3]}$ coincides with the average beginning of CO₂ uptake by coniferous trees (Kolari et al., 2007; Wallin et al., 2013)" I didn't know that evergreens only uptake CO₂ starting in May. I was convinced they always did. Is it true? Doesn't that contradict what you wrote in line 309 ("We base our temperature acclimation of coniferous trees on experimental results on Norway spruce which were found to be active already at rather low air temperatures and can reach 60% photosynthetic activity as early as day 100 (Kolari et al., 2007; Wallin et al., 2013).")? Here you state that photosynthesis is already active at DOY 100 and is at 60% of its maximum! It is true that coniferous can, in principle, take up CO₂ also in winter and the fact that they can suffer from frost drought shows that. With respect to the spring peak that was mentioned earlier, this means that coniferous trees may experience these elevated ozone concentrations regularly. Concerning the contradiction in the statements – if the conditions are right, coniferous trees can quickly reach 60% of their max photosynthesis. Conditions to keep photosynthesis up would occur from day 100 onwards. To make sure that we do not falsely accumulate too much ozone from the spring peak periode (when coniferous trees show on-off photosynthesis), we consulted satellite observations to extrapolate the dates when vegetations' photosynthesis (including coniferous trees) at Svanhovd rose over the detectable threshold.
- Line 233. "In July–September (day 182–273), ozone is occasionally almost completely depleted. This hints to ozone uptake by vegetation even at low light intensities during midnight sun conditions in combination with stable planetary boundary layer conditions preventing mixing of ozone rich air." I don't understand the con-

nection. What does the night uptake have to do with the occasionally complete ozone depletion? This was actually a working hypothesis and should be removed from the text.

- Line 244. "if a normal distribution is assumed" Are you sure that the distribution is normal and not lognormal or something else? There are some literature on the type of statistical distributions for variable such as Temperature, Rain, etc ... Moreover, looking at your Figure 6b the distribution of the irradiance seems to be a Poisson distribution. The data shown in Fig.6 are not the same as used to calculate the results for Fig.5 because the data selection is different. But we should of course check and describe the underlying distribution of the data selection used in Fig.5.
- Figure 5, caption. "dashed lines indicates statistical significance" Statistical significance of what? By means of what test was it obtained, at what alpha level? And what are the numbers on the top right of each graph? Regarding the test, see response to "L205" above. The number at the top right are described in the caption: "The annual positive/negative deviations are indicated in the respective corners (right upper/lower)." This is the percentage of days on an annual basis that exceed the 1σ level. This indicates the overall deviation from a normal year. Only temperatures in 2018 were abnormally high compared to the climatology on the 1σ level.
- Line 251. "deviated significantly from the climatology on the 1 sigma level." Here the standard deviation is used as reference for the significance. But the significance of the deviation should be statistically tested in another way. Regarding the test, see response to "L205" above. Our hypothesis test is that you would expect $\approx 18\%$ of days exceeding the 1σ level (normal distribution). If we observe a higher percentage of days, we find an exceedence from the expectation. If this is not proper in terms of statistical terminology we can remove the term "significant" and call it exceedence (in case this is deemed a better term).
- Line 262. "We use the bias-corrected and cross-calibrated ozone climatology (Falk et al., 2021) and assess the monthly significance of the ozone concentration anomalies in 2018/19." "Bias-corrected cross-calibrated" ozone? What is it? And what is the "significance" of the concentration anomalies? Please explain. Please see response to "Lines 206-208.". Cross-calibrated means that we used data from other monitoring stations in the region to reconstruct the missing data at Svanvik. Regarding "significance", see explanation above.
- Line 267. "Further, we presume that fVPD and fSWP suit our vegetation types and no acclimation is necessary for these." This statement is questionable, because in cold conditions VPD can be high (you also told it in the conclusions) and the water in the soil can be limiting because partially unavailable due to freezing or other. The referee is correct that VPD can be high due to cold conditions and thus become a limiting factor. We did not find any unexpected or abnormal

behavior of stomatal conductance with respect to VPD which indicated that the VPD limitation on stomatal conductance might be more general applicable for high VPD caused by either drought or cold. We are not stating that the response to VPD or SWP could differ but in subarctic vegetation we have to assume that they are fair enough. Water availability due to frozen soil is most likely not represented in the DO₃SE model and would be worthwhile investigating further. (Lisa?)

- Line 272. "but a substantially higher number of peak [O₃] were observed in 2018 than in 2019." How can you tell it if O₃ measurements for all months of March, April and July are missing in 2019? I don't seem to see any differences between 2018 and 2019 As we have indicated, the measurement down time in March and April was planned because most vegetation (in particular the one in the ozone garden) was covered by snow and hence not photosynthetically active (as the referee has noted, too in their comment regarding BVOC emission). "but a substantially higher number of peak [O₃] were observed in 2018 than in 2019." refers back to Sect. 2 L188-195. Peak concentrations were calculated as exceedence over 40 ppb in summer, explicitly excluding the spring peak. Even with two weeks of data missing July 2018, we found 50 elevated ozone event in 2018 compared to 18 times in 2019. This should qualify as "substantially higher number".
- Line 291 " Note, however, that these parameterizations are hypothetical and have yet to be verified by experiments." Figure 6a. Looking at the graph I understand that you assume an adaptation of the subarctic grasslands to the temperature distribution of the last decade (climate already changed) and not to the historical temperature distribution at your site. Is it reasonable to hypothesize such a rapid adaptation of vegetation to the new climate conditions? Indeed it is not reasonable for long-lived species such as trees or coniferous dwarf scrubs and depends on the ecological turn-over periode. Competition would lead to a decrease in number of individuals of those species which natural acclimation range differs strongly from the new norm. If we look at a grassland as sum of all species the whole grassland may rapidly adapt. (Does this make sense Ane? Would Vigdis' results back this up?)
- Line 298. "We construct cold as representative for a species that is more tolerant to cold temperatures, but slightly less efficient at warm temperatures compared to MM. This is accomplished by moving T_{opt} towards cooler temperatures while keeping the other parameters fixed to MM values". From Figure 6a and Table 1 I see that for the "cold" parameterization not only T_{opt} was moved, but also T_{min} for (e.g. for grassland). Indeed, we will correct the text.
- Figure 7. The g_{stom}/g_{max} ratio in the subarctic parameterization with PPFD0.8 is greater in the morning than at noon. How then the choice of PPFD0.8 is explained? Please comment on this in the text.
- Line 334. I don't understand how we can say that the differences are "substantial". I don't see much difference between deciduous trees (a) and grassland (c), sorry.

- Line 336. Using net photosynthesis to calculate leaf emergence is not completely justified. Leaves are likely to be present and active well before gross photosynthesis equals heterotrophic respiration (eg. soil respiration). Gross photosynthesis should be used to calculate Astart and Aend instead.
- Line 354. "A sample of downy birch leaves collected at Svanhovd had an average length of (3.0 ± 0.5) cm" Were top-canopy leaves sampled? How many leaves were collected to get ± 0.5 cm standard error?
- Line 355. "We used 13.5m height" Why was this value chosen? What is the meaning of a height between the average tree height and the maximum tree height? Perhaps it would have been more reasonable to use the average height.
- Line 360 and following. POD1 was calculated by gap filling the data, right? Because there is a lot of data missing in the middle of the season. Or were POD1 compensated for missing data? If so, how? Please confirm it by writing it in the text.
- Line 369. "Due to the shape of flight, a symmetric variation". Symmetric variation of what?
- Line 370. "We find that an opening of stomata at lower light intensities can cause higher sensitivity to drought conditions." Please, explain where we can see this. Graph 8 is not clear at all to me. And then, "sensitivity" of what? Of plants? Of POD1?
- Line 373. "The magnitude of these effects varies between PFTs as well as years, but the predicted ozone uptake for the bespoke temperature parameterization is always larger than for the MM parameterizations and of the same order of magnitude as the variability between the years studied here." What effects? "Of the same order of magnitude as the interannual variability...": can you conclude it by comparing only two years? The same for line 402
- Table 4. Have the percentage of reduction been calculated taking into account pre-industrial concentrations as prescribed by the MM? What are meaning of the superscripts? And, above all, why some superscripts indicate a range (e.g. 1.9 ... 2.1)? I did not understand how the stdev of the MM estimation was calculated, sorry.
- Line 416. "we have developed bespoke parameterizations" it seems a bit strong statement to me, you have not developed any new tailored parameterization, you have only hypothesized one. There is no one experiment nor comparison with experimental results in your work.
- Line 417. "The comparison between meteorological conditions in 2018 and 2019 and their divergence from climatology allowed us to assess the influence of key environmental variables such as temperature, PPFD, and precipitation on vegetation

susceptibility to O₃ damage in light of future changes as may occur under climate change” I did not understand where all this ”divergence with the climatological average” of these two years alone lies, sorry.

- Line 432. ”With respect to ongoing climate change, a clear positive trend emerged in length (5.2d decade⁻¹) of the growing season that is almost equally distributed between earlier start (2.9 days decade⁻¹) and later end (2.3d decade⁻¹) (Appendix Fig. A1).” How did you figure it out? Have you been doing retrospective MODIS analysis for 30 years? Or do you have a publication to quote?
- Line 435 and following. ”visible damage” Visible damage and POD can be totally unrelated, as demonstrated by some research conducted on agricultural species. I recommend caution in stating that the O₃ peaks causing the visible symptoms can result in a biomass reduction (damage).
- Line 441. Does ”damage” mean ”visible leaf symptoms”? Or does it mean biomass reduction?
- Line 456. ”We found that soil water potential under 2018/19 meteorological conditions was negligible” What does it mean? That there was no water in the soil (SWP were negligible) or that the effect on the POD of the presence or absence of SWP in the calculation was negligible? Please clarify.
- Line 461. ”better suited” Point 1 is questionable. Also point 2 is questionable. How can you say that the MM parameterization does not capture the plant physiology of subarctic vegetation if no comparisons with physiological measurements taken on subarctic vegetation are presented?
- Line 469. ”However, the decline of this ozone spring peak is partly caused by the uptake of vegetation” Are you sure? Please cite a reference.
- Line 491. ”Automation of the here proposed PDF-based acclimation using machine learning techniques could overcome these issues in the future” What does it mean? Please explain. Make an example.
- Figure A1. How was the length of the growing seasons in the various years identified? By satellite? Other method? What does the gray band represent?
- Line 511. ”with f_{min} , D_{min} , D_{max} describing the relative stomatal conductance to changes in vapor pressure deficit.” It is not clear. Please, clarify what D and f_{min} are, and their units.
- Line 517. ”The DO3SE model as described in B  ker et al. (2012) is used to simulate SWP₀ across a PFT specific root depth according to the Penman–Monteith energy balance method that drives water cycling through the soil–plant–atmosphere system” I cannot understand how the P–M energy balance is used in DO3SE to derive the SWP. Please explain in detail.

- Line 525. "the concentration at the upper surface of the laminar layer for a sunlit upper canopy leaf" At what height was the O₃ concentration measured? If it was not measured at the top of the canopy (10m for trees or 10 cm for grassland), how was the O₃ concentration at the top canopy calculated? Please explain in detail.
- Line 526. What does r_c represent? Is it the cuticular resistance or the bulk canopy resistance? What is its value?
- Line 528. Can you explain where that formula for calculating the flux comes from? Why is there $u(z_1)$ in? And what is the z_1 height?
- Line 531. What is the z_1 height? Where is it?
- Line 535. Wind speed at 2 m: what is it used for? Please explain
- Line 550. Please explicitly describe the method used to gap-fill O₃ concentrations because it could be crucial.
- Section B1. The description of fPHEN is missing. Please, provide it. Again, how do you calculate the day-to-day SWP on your site? Please describe it in detail.
- Line 554. "From Fig. B1f) it is apparent that the mapping manual parameterized grassland would not have been able to grow in 2019." It does not seem to me that g_{stom} has been reseted at all. If this is the case, the premises of the work appear weak.