GENERAL COMMENTS

The manuscript addressed a relevant issue related to the ozone risk assessment in the Subarctic. However, the presented study is highly speculative and is not based on any experimental evidence (i.e. the parametrization of gstom is hypothetical and not obtained from any measurements made in the area). The justification of some methodological choices is weak, for example the need to modify the gstom parameterization because in 2019 the gstom predicted with the MM parameterization would have been too low. Many claims sound superficial to me, such as the one about the alleged deviation of the years 2018 and 2019 from the site's climatology, and many statements are inaccurate (see the specific comments). The manuscript is too long, and the organization is confused. The introduction is nice, but it’s too long and contains too many digressions to get to the point. The results are mixed in the methodology and there is no clear division between the methodological section and the results section (a purely results section does not exist at all). There are several repetitions, for example the section “Discussion and Conclusions” is over half a summary of the previous sections. The Appendixes contain inaccuracies and omissions. The English should be carefully revised because some sentences are ambiguous. In my humble opinion, though, the work is generally weak and the manuscript is not mature enough to be accepted for publication.

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.

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?

Line 77. “(<\text{O}_3> = 36-54\text{ppb})”
Please explain the formalism. What do different brackets mean?

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”

Figure 1. This figure was never referred in the text.

Figure 2. It does not seem to me that the O3 concentrations of 2019 are different from those of 2018. The spring peak could even be identical (although unknown, because in 2019 O3 measurements started about 20 days after the spring peak)

Line 146. Please make clear the acronym PFT on first use

Line 165. “luftkvalitet.no”
What is it? And EBAS? Please make them clear.

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.

Line 190. “high ozone concentrations ([O3] > 40ppb”
It is strange to read that O3 concentrations above 40 ppb are “high” concentrations.
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 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?

I do not understand. Please, explicit the methodology.

“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.

“Darker colors indicate higher probability to observe these values.”
“On top of the density distributions, a 10 days average of daily mean (h[O3]i10d) is displayed together with 1sigma 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?

“The decline in h[O3]i 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 doy 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!

“In July–September (doy 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 connection. What does the night uptake have to do with the occasionally complete ozone depletion?

“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.

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?

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.

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.

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.

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

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?

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 Topt 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 Topt was moved, but also Tmin for (e.g. for grassland).

Figure 7. The gstim/gmax ratio in the subarctic parameterization with PPFD0.8 is greater in the morning than at noon. How then the choice of PPFD08 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 fmin, Dmin, Dmax describing the relative stomatal conductance to changes in vapor pressure deficit.”
It is not clear. Please, clarify what D and fmin are, and their units.

Line 517. “The DO3SE model as described in Büker et al. (2012) is used to simulate SWP 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 rc 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 gstom has been reseted at all. If this is the case, the premises of the work appear weak.