Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-322-AC4, 2020

Biogeosciences Discussions

Interactive comment on "The rising productivity of alpine grassland under warming, drought and N-deposition treatments" *by* Matthias Volk et al.

Matthias Volk et al.

matthias.volk@agroscope.admin.ch

Received and published: 17 November 2020

Below, please find the authors point by point replies. For ease of reading we quote the comments with a '?' first and start our responses with '>'

? General Comments In their manuscript titled "The rising productivity of alpine grassland under warming, drought, and, N-deposition treatments", the others describe a novel experiment in which monoliths of soil and turf were transplanted across an elevational gradient combine with fertilization and water addition treatments. After four years of growth in the transplanted location, the others describe how plant productivity in the monoliths responded to the interaction of different temperatures (comparing climate at the transplant location to the original site where the turfs were harvested

from), fertilization, and increased moisture, as well as the interactive effects of these three treatments. The results of this study showed that intermediate levels of warming increased plant productivity, even in drier conditions. Increasing the precipitation received by some monoliths had only marginal effects on plant productivity, while fertilizing the plots with nitrogen solutions had no discernable effect on plant productivity. While this experiment is truly novel in its use of monolith transplants to simulate climate change in conjunction with two additional global change treatments in order to understand how multiple facets of global change will impact productivity, I have several concerns regarding the framing of these treatments, the metrics used to communicate and aggregate results, and the overall clarity of the manuscript. In particular, while transplanting monoliths to new elevations does of course impact climate, and in some cases results in warming, characterizing this experiment as a "warming experiment" is disingenuous. I encourage the authors to refer to their experiment as is, a transplant experiment across an elevational gradient. Furthermore, it is also a misnomer to refer to the precipitation manipulation component of this experiment as a "drought treatment", as water was added to some monoliths instead of removing precipitation, as when using rain-out shelters etc., to simulate drought.

> Indeed, in the headline we imply that we have a warming treatment, even though what we apply is an altitudinal transplantation treatment. Analogously, drought, as a productivity limiting factor, is not a treatment, but also a consequence of the downward transplantation in our experiment. We chose this wording to quickly convey motivation and relevant drivers the experiment. The supplementary precipitation (not mentioned in the headline) is a treatment to mitigate drought conditions and addressed appropriately in the text.

? My detailed line comments below elaborate on these concerns as well as my suggestions and critique of the metrics that the authors chose to describe climate in this study.

? Line Comments 34–"... to have beneficial effects": Beneficial effects on what?

> Sorry to be unclear. Old sentence reads ‘First, mitigation of the thermal growth limitation is likely to have beneficial effects.’ New sentence reads ‘First, mitigation of the thermal growth limitation is likely to have beneficial effects on productivity.’

? 35-36: Clarify what you mean by “initial water supply”... Water resources at the beginning of the growing season are generally plentiful? But this would be the case only for plants that emerge early in the growing season, i.e. depends on phenology of plant species.

> Yes, water resources at the beginning of the season are generally plentiful. As we state behind the comma ‘. . ., because even a small winter snowpack supplies a large soil moisture resource in spring.’ Plants in subalpine grasslands are all perennial and usually start greening even before the snow-cover has completely disappeared.

? 38–“kg N ha-1 a-1”: These units are unconventional, instead of a-1 (per annum?) I typically see yr-1 when describing nitrogen deposition rates.

> Yes, ‘per annum’. Not unconventional. The SI convention for English year is ‘a’

? 45–“...showed a twofold productivity increase”: In response to what treatment?

> In response to increased summer temperature. The whole sentence in the manuscript reads ‘For example, tundra vegetation showed a twofold productivity increase, driven by increased summer temperature (Van der Wal and Stien, 2014)’. For the revised MS we will complement ‘ . . . up to twofold . . .’ to better reflect the quoted authors statement.

? 47–“...grasses were favored over forbs and sedges by drought and warmth”: This seems unclear, what do you mean by "favored by drought and warmth"? Productivity of forbs and sedges increases with warming and drought?

> Sorry, unclear. By ‘favored’ we mean that grass relative abundance increased at the expense of sedges and forbs. New sentence reads ‘In contrast, Liu et al. (2018) combined long-term observations with a manipulative experiment to find that total net

primary productivity (NPP) in Tibetan grassland remained unaffected, though relative abundance of grasses was increased at the expense of forbs and sedges by drought and warmth.’

? 61–“...if only a short or linear segment out of a larger range of biologically possible responses is represented in the data.”: There is some indication that productivity relationships revealed in manipulative experiments actually encompasses even more variation than occurs naturally (see Jochum et al. 2020. Nature Ecology and Evolution).

> Here we are making a point to include many factors and factor levels in climate change experiments, in order to avoid wrong interpretations when interpolating between data points. The original sentence reads ‘Not only can a low number of treatment factors, but also a low number of treatment levels invite overly simplistic interpretation of experimental results, if only a short or linear segment out of a larger range of biologically possible responses is represented in the data.’

We do not understand how this is related to your comment above. With respect to biodiversity experiments the Jochum et al. paper finds that biodiversity experiments ‘have greater variance in their compositional features than their real-world counterparts’. But based on their analysis they later conclude that this does not impair the applicability of the results of biodiversity experiments: ‘ . . . our results demonstrate that the results of biodiversity experiments are largely insensitive to the exclusion of unrealistic communities and that the conclusions drawn from biodiversity experiments are generally robust.’

? 67–I think that I am still confused by what you mean by "factor levels"... Does this refer to consideration of multiple global change factors, or does it refer to the magnitude of the global change treatment imposed by the experiments?

> Indeed, factor levels, as in our sentence ‘. . . the outcome . . . depends to a large degree on the chosen factor levels . . .’ refers to the magnitude of the chosen global

change treatment factor, e.g. an N-deposition with levels of 0 (control), 3 and 15 kg N ha-1 a-1 on top of the background deposition.

? 68– “Here, we present four-years of treatment results from a field experiment in the Swiss Alps.”: This statement is an important introduction of your experiment, and as such, you should be more descriptive than "treatment results from a field experiment". What types of treatments specifically were involved in your field experiment, and were any of these treatments applied simultaneously to study interactive effects?

> As reviewer #2 demands, the lines 68-75, following the sentence quoted in l. 68, wrap up what we did in the experiment. Only it is not in the first, introductory sentence.

? 83–“monoliths (ML)”: I do not feel that it is necessary to use an acronym for one word, and stating monolith regularly instead of ML will improve the clarity of your manuscript.

> We agree that abbreviations should be used conservatively, but we are undecided about this issue. After all the term occurs 36 times across the MS.

? 102-103: This sentence is rather unclear. What do you mean by standardizing harvests and the "zero-year" and "acclimation" distinctions? This aspect of your methods deserves an elaboration.

> We recognize the cause of confusion. Indeed, the distinctions ‘zero-year’ and ‘acclimation’ are obsolete in this place. They derive from the chronology of establishing the experiment. The ‘standardizing’ harvests in these first two years served to homogenize the canopy of the monoliths, that were originally grazed and therefore had more heterogeneous canopies than mown grassland. New sentence: ‘Standardizing harvests were done in 2012 and 2013, to homogenize the canopy of the previously grazed monoliths, that had more heterogeneous canopies than mown grassland.’

? 111-115: I find your naming convention, using the ’CS’ designations, to be needlessly confusing. These are simply sites along an elevational gradient, so why not refer to them either by their numeric elevation (i.e. 2360 m) or simply as Elevation 1 (lowest

elevation), Elevation 2.... etc., instead of introducing a less intuitive naming system.

> We chose the term ‘climate scenario’ (CS) to make clear, that these sites are associated with a very complex treatment, containing a number of factors. Namely, the treatment includes changes in soil moisture, temperature and growing season length. When the data is presented in the text or in figures over x-axes, that designate temperature or moisture values, it is more intuitive to use names like ‘climate’, that commonly associated with temperature or soil moisture values. But similar to the monolith ML issue we work on avoiding the abbreviation CS.

? 116–“...6 CS, 6 MLs from each of the six sites of origin”: I find your naming convention, using the ’CS’ designations, to be needlessly confusing. These are simply sites along an elevational gradient, so why not refer to them either by their numeric elevation (i.e. 2360 m) or simply as Elevation 1, Elevation 2.... etc., instead of introducing a less intuitive naming system.

> Please compare the response above (l. 111-115)

? 119–“...were filled with soil to prevent air flow”: Where did this soil come from? Bulk soil from each specific elevation/origin location?

> The soil used originates from the respective scenario site, i.e. from the pit that was dug to accommodate the transplanted monoliths. This means the ‘filling-soil’ was not the same as in the monoliths (that come from six different origins). This does not affect the individual turf monoliths soil properties, because the monoliths remained in their drained containers for the whole duration of the experiment, so that the monolith-soil was isolated both from neighboring monoliths with different soil and from the ‘filling-soil’.

? 121–“cross-factorial design”: Full-factorial design? I’m unfamiliar with "cross-factorial" experimental designs.

> Yes, thanks, we got that wrong. We change that to ‘full factorial’

? 153: This sentence is rather unclear... Temperatures were summed across one day?

> Unfortunately we can't find a reference to temperatures in l. 153. Our best guess is, your comment refers to l. 148 ff, saying 'The thermal energy was expressed as degree day values (DD0°C), resulting from hourly air temperature means above a threshold of 0 °C, added for one day, then divided by 24.'

We are sorry we were not clearly describing our standard procedure to calculate degree days from hourly temperature means. Indeed, there is a plethora of 'degree days', tailored to suit many specific purposes and there is no single convention. We improved the situation and the complete section now reads: 'The available thermal energy was expressed as degree days (DD) above a threshold of 0 °C (DD0°C). To derive DD we calculated the sum of hourly temperature means above 0°C during one day, then we divided this sum by 24 hours. To quantify the total thermal energy available for growth, we summed degree days during the snow-free period between the annual harvests (DD0°Ctotal), considering that the perennial vegetation continues to grow after mowing.'

? 154-156: This threshold seems particularly arbitrary, and I think that the use of a threshold in general is not necessary here. Why not simply present the mean growing season soil volumetric water content for each site/each season? This metric is much simpler and more intuitive for readers to understand and compare your results across the elevational gradient.

> We considered using mean growing season soil volumetric water content and dismissed the idea. The reason is similar to the problems arising when using mean temperatures: The plants do not experience 'mean' water contents, when coping with environmental growth limitations. For example, when plants experience a wet month after a dry month, the mean soil moisture may suggest perfect growing conditions, when they were bad indeed. We do not think that an increasing number of dry situations is less intuitive than a decreasing number of soil water content.

? 161-162: Why does the amount of precipitation added to each monolith vary between years?

> The application of the irrigation treatment was limited by the concurrent availability of workforce and occurrence of dry soil situations. We would have preferred to add more water, but did not manage to.

? 168: Listing the chemical formula of ammonium nitrate is not necessary.

> Thanks for mentioning that.

? 226: Is there some type of relationship between atmospheric N-deposition rates and elevation? Perhaps describe N-deposition rates across the entire gradient, not just at the middle and low points of your elevational gradient.

> We only have data for the second highest site CS2reference (3.3 kg N ha-1 a-1) and the lowest site CS6 (4.3 kg N ha-1 a-1). This difference likely reflects the distance of the CS from the (agricultural) N-sources. CS6 (1680 m a.s.l.) is close to the village, CS2reference (2170 m a.s.l.) is further up the mountain.

? 236: What does non-continuous mean? Non-linear?

> We wrote 'We observed a small, non-continuous increase of precipitation with altitude during April – October. The recorded annual precipitation sum was somewhat larger than the sum for the growing period (Tab. 2).' We meant to say that precipitation was not continuously rising with altitude. The second sentence refers the reader to Tab. 2, that contains the precipitation data for all sites. Also, we did not mean 'non-linear', as we did not attempt to fit a (non-linear) model to the data.

? 239–"...only one third of the pre-harvest period was dry": It is definitely a misnomer to describe conditions of lower than 40% moisture content as "dry". In fact, in most alpine systems, 30% moisture content is considered ideal moisture conditions for optimal microbial activity (see Hawkes et al. 2017 PNAS for a relevant discussion related to respiration and soil moisture). I would highly suggest re-characterizing the way in

which you describe soil moisture in this manuscript. Instead of creating a binomial in soil moisture conditions around an arbitrary 40% moisture content threshold, why not just describe average soil moisture across the growing season on a continuous scale, i.e. just state average growing season soil moisture for the pre-harvest period.

> We agree that it would be advantageous to find a better term than plain 'dry' for sentences like this. This will be improved throughout the revised manuscript. As explained in the Material and Methods section (l. 153-156), the 40% threshold was neither chosen arbitrarily nor do we imply values < 40 % to cause drought stress. Instead, the threshold was found empirically to provide a good contrast between CS and years. It is also clear, that the soil water availability is decreased in periods with an increased percentage of days with SWC < 40%. We find the Hawkes et al. 2017 paper brilliantly studies the legacy of local climatic history on differential, local microbial adaptation. They find that microbial respiration is effectively locally specialized to soil moisture conditions. We could not find references to plants, plant productivity, ideal moisture conditions or alpine sites. We do not agree with the idea to describe the water related growing conditions as 'average soil moisture across the growing season' and reiterate our response to a comment above (154-156:): 'We considered using mean growing season soil volumetric water content and dismissed the idea. The reason is similar to the problems arising when using mean temperatures: The plants do not experience mean water contents, when coping with environmental growth limitations. For example, when plants experience a wet month after a dry month, the mean soil moisture may suggest perfect growing conditions, when they were bad indeed.'

? 248-249: Because you describe soil moisture conditions in the previous section using percent dry days, we have no way of understanding how this transplantation effect on soil moisture conditions (described using VWC) might interact with your other treatments.

> Yes, there is a way of understanding the transplantation effect. In the section quoted, we state both the SWC for transplanted monoliths and the undisturbed grassland in

the simplifying 'average SWC' metric. In any case, the good message will be clear, because the difference stated is only 1% vol SWC.

? 251: I would suggest that productivity is the more appropriate term, consistent with literature in this area of ecological research, to describe your response variable.

> We will change that to 'productivity'. Unfortunately we can only offer a crude proxy for 'productivity' (net ecosystem productivity). The harvestable part of the canopy is less than net primary production. Also, due to the fixed cutting height, this metric creates an overestimate of positive effect sizes, that is larger when the yields are small. For the above reasons we replace the unspecific term 'productivity' with 'yield' if we can.

? 259: In order to show evidence to support this claim, I would like to see a figure and the related statistics that shows the relationship between the productivity effect size (productivity in transplanted monoliths - productivity in control monoliths that were reinstalled at the same site / standard deviation of productivity across all monoliths) regressed against the temperature difference from the monolith's original climate and the transplanted climate. In other words, how much of the change in productivity is explained by change in temperature following transplantation?

> Strictly speaking, we claim ' . . . we found a highly significant effect of the CS . . . ' in l. 259. Here, we do not claim that temperature caused the significant differences in yield. Instead, we refer to the climate scenario (CS) as a whole, because it is one of the strengths of our experiment, that we simulate climate change in the mountains as complex climate scenarios, including simultaneous changes in thermal energy, growing period length, water availability and increased pollutant deposition. An abundance of aspects of the yield response is displayed in four panels (Fig. 1 and Fig. 2) and in Tab. 4, with the statistics shown in Tab. 3 (+ Appendix Tab. A2 and A3). But, as demanded by the reviewer, we have also broken down our analysis (generalized additive models) to individual, environmental parameters of CS, namely degree days (DD0°C) and < 40% SWC conditions (dry days %). This information can be found in Results l. 299-

302 (DD0°C) and l. 303-308 (dry days %) and in Mat. & Meth. l. 215-219.

? 260-261: What does "intermediate warming" mean here? Describing this result as "monoliths that experienced X-Y degrees of warming by being transplanted to warmer climates at lower elevations relative to climate at their original location showed increases in productivity".

> 'Intermediate warming', is a term we use in the context of distinguishing the climate scenario sites (CS). It refers to those CS where the altitude related warming component of the climate change scenario is in the middle between minimum warming and maximum warming. We think that once the reader arrives at the Results section, 'intermediate warming' in the context of this study is clear enough. We aimed at keeping the Results section comprehensive, but short. Thus, we would rather not repeat the Mat. & Meth. as suggested.

? 262-264: This sentence is confusing. 2016 was the year in which productivity, on average, was highest, but this was only the case at two sites? These two statements seem to contradict one another.

> Sorry for causing confusion. Actually, we don't say that it was only the case at two sites. Only CS5 did not show maximum yield in 2016 (the corresponding numbers are in Tab. 4). Our l. 262-264 says 'In the year of the overall maximum productivity (2016), both the coldest site CS1 and the warmest site CS6 produced their respective record yield (Tab. 4).' We will replace 'both' with 'also' to be more clear. We use the term 'also' to draw the attention to a counterintuitive situation: Despite transplantation into contrasting environments (cooler at CS1 and substantially warmer at CS6), production of the maximum yield coincided with the weather conditions of the same year.

? 298: The title of this section seems to not relate to the results described within the section. You already stated that each elevational site is characterized by different temperature and precipitation regimes in your methods and in previous sections of the results. Should this section describe the relationship between productivity and climate

at each elevation?

> Very helpful! This section really describes the relationship between biomass yield and those environmental parameters (warmth and moisture) that we quantified for the individual climate scenario (CS) sites. This is different from the approach that treats CS as categories that integrate multiple climate change aspects. Accordingly, we will change the title. We suggest '3.2.5 Biomass yield response strongly relates to temperature and soil moisture changes'

? 325-326: Are there examples of other papers whose conclusions about the use of degree days instead of mean temperatures over the same time frame?

> Particularly in environments with strong temperature contrasts (day/night, summer/winter) like mountains or deserts, the use of DD does constitute a much more valuable metric for plant usable thermal energy. Similar to mean soil moisture values, mean temperatures can be extremely misleading, because a sequence of hot and freezing temperatures may well result in a comfortable average temperature that the plant has never experienced. Indeed, the whole concept of mean values is of quite limited use, just like dressing for the outdoors according to the current calendar month, instead of testing the air in front of your door. Some examples for the use of DDs in the context of grassland research are - Dukes et al. PLoS Biology 2005 (Jasper Ridge Experiment (CA)) - Fridley et al. Nature Climate Change 2016 (plant funct. strategies of 20 years UK grassland warming) - Wang et al. Ecology Letters 2020 (extremely dry Tibetan alpine grassland) - Wilsey et al. Journal of Applied Ecology 2018 (42 US grassland sites) - Zimmermann and Kienast Journal of Vegetation Science 1999 (Swiss alpine grasslands)

? 333-341: This section would benefit from a description of why the authors suspect that warming beyond "intermediate warming" was not associated with the same boost in productivity that was associated with intermediate warming.

> Good point. Originally we only implicitly described that (l. 338-341). New formulation:

‘In the extreme treatment at lowest CS6 (+3 °C Apr.-Oct., +2.4 °C annual mean) the positive response to warming finally ceased to increase, but yield was still somewhat larger than at CS2reference. This demonstrates that the growth benefit from the larger thermal resource compensated for a radically smaller soil water resource (compare Figure 2 A & B). But the comparatively low growth response suggests, that the water supply at CS6 has already reached a critically low level.’

? 337–“cockchafer (Melolontha melolonth) infestation: Please describe what this organism is and how it is relevant to variability in productivity.

> The Cockchafer is a bug, its larvae feed on roots. When there are many, they may kill the vegetation. The Cockchafer (together with the Locust) is probably the one insect that is best known to the public for its periodical mass flight-years. In these years it is a major pest. We would prefer not to add too much general biology to the text.

? 347-349: Grammatical errors and diction in this sentence make it unclear.

> Reformulated sentence to be clearer: ‘Also, the dramatic temperature dynamics that occurred during the past 12,000 years of the Holocene interglacial, suggest that temperature adaptations, that are still contained in modern plant genotypes, may actually match not only today’s weather, but also warmer and cooler climate conditions.’

? 358: I think this statement describes my point about eliminating your use of the "percent dry days" metric entirely... Your results, using this metric, prevent readers from relating the soil moisture conditions present in your experiment to soil moisture conditions elsewhere. Furthermore, describing soil moisture conditions less than 40% as "dry" is a misnomer.

> We admit that between-experiment comparisons of soil moisture conditions, or rather the water availability for plants, is close to impossible. The reason is that A) different plants have different capacities to exploit the moisture resource. That means that a species from one experiment thrives at the same SWC when a species from another

experiment dies. B) different soils have different water potentials (osmotic plus matrix potential). As a result, soils with the same vol. % SWC may have totally different water availabilities from a plant perspective. Consequently, we chose to generate a wide range of within-experiment soil moisture conditions for comparison, rather than refer to literature values. Else, we believe that quantifying environmental conditions by describing them as more or less dry days occurring, gives the reader a perfect idea of which situation was more beneficial and which was less so. As stated above (reviewers comment l. 239) we agree that ‘dry days’ is not a perfect choice, as dry implies a critical situation for the plant. But the metric ‘% days < 40% SWC’ serves very well to distinguish between scenario sites, years and irrigation treatments. This highlights its value for describing the situation. Please cf. tables 2 and 4.

? 380: What caused increased evapotranspiration at CS5? Is it possible that too much rainfall, either ambient or added as part of your irrigation treatment, could cause leaching of important soil nutrients, with higher VWC leading to lower productivity? This might be especially relevant in monoliths that received both an irrigation and fertilization treatment.

> We considered that it was likely the higher temperatures in those climate scenario sites (CS3, CS4, CS5 and CS6), that were located downslope from our reference site CS2reference, that caused higher evapotranspiration. We have no reason to assume that there was too much rain. The nearby federal meteorology station recorded 662 mm/year during the experiment, while the 1981-2010 mean is 706 mm/year. Indeed, the Massenerhebung effect creates a continentality of the climate that makes inner-alpine valleys like the Engadin quite dry. Please also compare tab. 2. Our irrigation treatment only added 12-21% of the seasonal rainfall, the nitrogen deposition treatment was equivalent to 20 mm precipitation per year for all monoliths. This is not a likely scenario for nutrient leaching.

? 399-402: These are the only lines of this section of your discussion that reference your results directly. These sentences should be moved up in this section, and you

should eliminate the references to other experiments with results that contradict what your experiment found, as this section is very unclear as curssrently written. Which of these citations and theories help explain your results? Remove the rest.

> The references to other N-deposition experiments are carefully chosen and reflect the best comparisons available. Early experiments applied high doses to test for N-limitation in general. The deposition rate in later experiments was lowered, to approximate realistic atmospheric N-deposition rates. Our experiment approached the critical N-load, the limit between a responding and a not responding ecosystem. In the section in question, the relevant differences compared to our experiment are highlighted: Sometimes the deposition rates were substantially higher, sometimes single species responses or general plant community changes were reported. Summarizing these contrasts we conclude that 'the cold-adapted, mature and low productivity grassland either responds with a >5 year time lag, or that the N-deposition treatment was below the critical load for aboveground biomass responses.' We feel this is a reasonable line of argument. Also, we consider the Discussion to be a place to reflect on the state of science in the field in general, as opposed to collecting arguments that 'help explain' our results.

? 426–"This implies that subalpine grassland productivity has likely not increased during the past century warming": This statement is in no way supported by your results.

> We found that those monoliths that were subjected to a cooling treatment (at CS1), such that they experienced the temperature conditions of the 1920s, did not show a reduced growth compared to the climate scenario at CS2reference with 'modern' temperatures. What else should we conclude from that, if not that the last 100 years of warming did not affect plant growth yet?

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-322, 2020.