In this study, seed count of *G. pneumonanthe* shoots was analyzed as a function of soil variables, shoot flower phenology, and species interactions. The aim was to study whether local soil variables directly, or indirectly, impacts plant reproductive output. This was done at one location and at one time point (in a period of ca. two months). Unfortunately, the study is seriously compromised by a poor introduction, flawed data collection, analysis and interpretation, and lack of novelty. My first biggest concern goes to the interpretation of seed count in one flower of a shoot as a proxy for plant reproductive performance, which it clearly is not. No effort was done to collect data on the plant level (instead of the shoot level), which is the level at which reproductive output should be measured, and the level that is relevant in terms of microclimate (or micro-environment). My second concern is that soil measurements at one time point are interpreted as microclimate. Microclimate is the product of temperature, precipitation, wind, sun, slope and many more local environmental factors, so it is inappropriate to use this terminology. This flaw, together with an inappropriate definition of plant reproductive output and biased statistical analysis, caused dramatic overselling. Please find below more details.

Title:

The title is extremely misleading. It involves the terminologies “microclimate” and “plant reproductive performance”, which were not studied.

Abstract:

Reproductive success is presented as the amount of seeds per flower, but surely one flower does not represent plant reproductive success. Without data on total reproductive success (across flowers and shoots), this study provides little contribution to our understanding of how micro-environmental variation impacts plant reproductive traits.

The abstract misses any clue to which methodology was used to perform this study. At least give some info about sample size and modeling.

The conclusion of the abstract is simply wrong: “These findings highlight that among-individual variation in small-scale environmental conditions within populations can cause variation in individual plant performance through multiple pathways.”. First, among-individual variation was not at all assessed. Assessments were on the shoot level, with serious pseudoreplication among shoots within plants. Second, individual plant performance was not assessed. Again, assessments were on the shoot level, and even only flower per shoot was assessed.

Introduction:

It is not clear why it is important to study reproductive success at microclimate scale. What differences are expected when studying this trait in populations from a microclimate vs. a macroclimate point of view? I mean, the exact same processes play here, no? Especially at this point, where it is unclear that the authors actually refer to soil variables, and not to climate. The authors need to emphasize the importance of studying soil variation in this context, because it likely is the only novel thing about the study (e.g. we already know that climate affects species’ interactions in *G. pneumonanthe*, we already know that climate affects flower phenology in *G. pneumonanthe*, we already know that species’ interactions affect reproductive output in *G. pneumonanthe*, ….).

Also, no effort was done to show what is already know about effects of micro-environment on these various processes in the species under study (*G. pneumonanthe*). That is a huge literature gap that needs to be filled, for which there is ample room because the introduction is extremely short.

It is not sure what are the implications of this study. You have some (yet biased) insights in flower seed counts as a function of soil data. What does this tell us exactly? Is it important for management?

Research questions 2 and 3 seem like exact copies of one another. Please make more clear. I think it should be something like (i) what are the relative contributions of direct vs indirect effects of microclimate on reproductive output, and (ii) how do plant phenology and species interactions moderate these effects?

Methods:

The number of reproductive shoots per plant was not reported, while clearly it is a key factor for the reproductive output of a plant.

Also, you assumed that individual shoots attract the butterflies, but it is the density of shoots that attracts the butterflies.

"eggs were counted on the whole shoots". Please clarify that this resulted into one egg count per shoot (because intuitively one expects per-flower egg data, which the authors don't seem to have).

How were the shoots marked?

For one shoot (and thus one plant), you counted the seeds from one intact and if available one preyed fruit, while (i) one gentian plant can have tens of flowers (even more than a hundred), (ii) the number of seeds per flower can vary dramatically between flowers within and between shoots of the same plant, and (iii) the proportion of intact vs preyed fruits can vary dramatically between shoots within and between plants. So there is no possible way in which such data can be representative for plant reproductive performance!

On top of that, you average the number of seeds across intact and preyed fruits, taking away any possibility of testing the impact of predation on seed output.

Shoot phenology was averaged in subplots (so please do not refer to "plant phenology" but to "shoot phenology" or simply "phenology"), and there is no information on reproductive output and predation at the plant level. It is ironic that the authors define micro-climate as the environment experienced by a plant, yet do not perform their study on the plant level.

This biased interpretation of plant reproductive output and phenology becomes even worse during statistical analysis, where pseudoreplication of shoots within plants was not accounted for, for example through a random "plant" effect (or by simply averaging within plants).

I am not familiar with SEM approaches, but I do now for a fact that any modeling approach requires validation (in addition to model fitting) in order to convince readers of their robustness and reliability.

Results:

You found an R^2 of 0.02 for the number of seeds in a shoot flower to be correlated with soil temperature. This likely is even overestimated given that there are problems of pseudoreplication. It is not a shoot that responds to micro-environmental variation; it is the plant that responds… So all shoots of the same plant will respond similarly. I also wonder: why not separating seed count of preyed flowers from seed count of intact flowers so that at least you can assess the impact of predation on seed count?

Also please realize that you try to explain multi-step processes (in the subsequent SEM analysis) for a nearly inexistent effect (given the R^2). So if you have a starting R^2 of 0.02, what is the meaning of R^2 in subsequent SEM analysis? Lets say a SEM R^2 of 0.5, is this 0.5 of 0.02 (so 0.01)? How do we need to interpret this result in light of the initial regression model?

Discussion:

The first section of the discussion already points to serious over-interpretation: No, your results do not show that the effects of microclimate on reproductive performance of the plant *G. pneumonanthe* were mainly indirect. They show that soil temperature (not microclimate) can, slightly (!), affect seeds in a flower (not reproductive performance!!) through predation. In addition, you found that *P. alcon* had a tiny (not strong!!!) effect on seed counts (not plant performance!!!). Etc…