Review of "Measurement report: Characterization of the vertical distribution of airborne Pinus pollen in the atmosphere with lidar-derived profiles: a modelling case study in the region of Barcelona, NE Spain" by Sicard et al.

The manuscript describes a first attempt to compare pollen concentration modelled with a 3D transport model with profiles retrieved from lidar observations. It is an interesting and timely topic as interest in bioaerosols is growing for both their impacts on health and climate. While the vertical distribution of pollen and other bioaerosols is relevant for their long-range transport and ability to impact cloud formation, it is not currently well studied. The paper is quite clearly written but could benefit from shortening. I generally support publishing the manuscript, but some revisions and clarifications are needed.

Major comments:

1. The way the references are used in the first paragraph of Introduction gives the false impression that the function and allergenicity of pollen were only recently discovered. Generally, papers that report the actual discovery should be cited, however, these facts have been known for long time and can be considered common knowledge and thus the relatively recent references to Moore and Weeb, 1983, Skjøth et al., 2013, Burczyk et al., 2004; Ellstrand, 1992; Ennos, 1994, D'Amato et al., 2007; Sofiev et al., 2017 that are attached to these statements should simply be removed.
2. Page 6, Equation (3): This equation seems to be incorrect. The authors claim to be using growth factors from Griffiths et al. (2012). Though it is not well described how they obtained the specific numbers, Griffiths et al. (2012) on their Figure 1 show average mass growth factors and the numbers presented by the authors in Table 3 look similar enough to that figure to assume it to be the source. With the assumption that the size of the pollen does not change with the water uptake as explained by the authors at the end of page 5, and given that the growth factors in Griffiths et al. (2012) are defined as the ratio of wet pollen mass to its dry mass ($\phi = \frac{m}{m_{dry}}$) the density should simply be the dry density times the growth factor:
$\rho_{pinus} = \frac{m}{V} = \frac{m_{dry} * \phi}{V} = \rho_{pinus-dry} * \phi$, as the volume V stays constant.

3. The 8 pollen types that Griffiths et al. (2012) average to get the growth factor do not include *pinus* pollen but do show significant variability in their water uptake in higher humidity conditions. Are the simulations expected to be sensitive to this value?
4. Authors have implemented a scheme based on Helbig et al. (2004) for pollen flux to the atmosphere. Equation (4) shows the flux in grid cell x at time t to be proportional to characteristic concentration P in canopy, defined by equation (5). Eq. (5) defines P as proportional to total seasonal pollen production from all trees in the grid cell and inversely proportional to tree height and area covered by pines. While this relation does provide pollen concentration in pine canopy, in its presented form it also implies that the flux on grid cell level is inversely proportional to the area covered by pines, e.g. a dense pine grove would emit much more than sparser forest with same number of pine trees. Is this corrected for in some other part of the model not described in the paper?

5. In the paper by Helbig et al. (2004) the emission model description mentions the step of reducing the total number of available pollens by the already emitted amount. Such step is indeed necessary to keep the cumulative flux from exceeding the total pollen production. How is this implemented in the current model?
6. Page 9, line 27: "we consider the first lidar measurement (225 m) to be a proxy of what it would be at ground level" - how correct is this for 59 um particles, especially in emission regions? Comparison with the findings of recent papers by Rojo *et al* (2019) and Hugg *et al* (2020) that investigate the near ground vertical profiles of pollen concentration based on pollen traps at different heights from ground could be helpful.
7. Page 9, line 29: “In the case of the model, we consider the first model layer, the center of which varies between 24 and 24.4 m over all the simulations considered here.” Is the model first layer 50 m thick also for 1 km resolution simulations? How well can a model with so thick layers represent vertical mixing in small scales?
8. Taking into account that the parameters in the formula for getting pollen concentration out of lidar observations (Eq.11) are all multiplied with each other could they be combined to a single unknown parameter (e.g. number extinction efficiency instead mass), which could simplify the presented sensitivity studies?
9. What are the assumptions behind getting *pinus* mass fraction from total pollen - what species, size and density is assumed for the other, non-*pinus* pollen and what are the uncertainties resulting from these assumptions?
10. Page 13, line 11 the authors write: " the emission source is widely distributed across the domain of interest (see Figure 1a). The latter allows a better analysis of the role of short-to-medium range transport and vertical mixing processes in the dispersion of this type of pollen." - How does the wide and uniform source distribution enhance the ability to analyze the role of transport and mixing?
11. "The model has significant skills to reproduce the temporal variability of the observations with sudden peaks on 30M and 31M or sustained high concentrations the midday of 28M or afternoon of 31M. " What is the p-value here? For instance, for daily set of 24 hourly observations, Pearson R needs to be > 0.4 to reach statistical significance at $p<0.05$, which does not seem to be the case according to Table 7.
12. Averaging FB over whole day including also very low concentrations should be avoided as it leads to misleading results that require paragraphs of explanation. It might make more sense to compute it separately for night and day.
13. Surface emitted heavy particles are likely to produce monotonously decreasing vertical profile most of the time, so good correlation between model and observed profiles is expected and sedimentation causing this profile to uniformly shift downwards should indeed not impact the correlation much.
14. Is the daily average correlation computed hour by hour and averaged or computed over the whole daily dataset?
15. It would be informative to see the lidar derived observations of pinus pollen concentration side by side with the model simulated ones in the lidar location as a 2D time-altitude plots.
16. It's not surprising that the highest resolution model doesn't reproduce the observed features for many reasons, including imperfect wind fields and emission depending on microclimate. I would suggest comparing the variability in model and observation (standard deviations ratio, maximum and minimum values or something similar) for investigating the added value of higher resolution model.
17. Page 24, line 15: “The Hirst observations are much more variable than the model concentration and the meteorology.” Hirst is a point measurement influenced by local emissions and thus it should be more variable than even 1km grid. Also, real trees are individuals that do not follow exactly the same

temporal flowering pattern and also release pollen when personally shaken by small scale turbulence. Concentrations higher up in the atmosphere are a mixture of upwind emission from many trees and thus expectedly smoother.

18. Assumption that mass fraction of pinus pollen would be constant through the whole boundary layer can be investigated using the model data by looking at the ratio of concentrations of pollens with different aerodynamic diameters simulated in the base case and enhanced runs.
19. The readability of the paper would benefit from shortening it. For instance, some detailed day by day descriptions of observations and model performance can be substantially shortened, given that the information is already present in the form of a figure or table.

Minor corrections:

1. The word “specie” is not the singular of “species”, in fact, “species” is both singular and plural. Please correct throughout the manuscript.
2. Table 5 and Fig 4 duplicate contain the same data, one can be removed
3. Table 8 and Fig 11 also duplicate each other
4. Page 2, line 27: ”The seminal works of Helbig et al. (2004) and Sofiev et al. (2006)” – The list of seminal works in pollen modelling should also include Schueler and Schlünzen (2006) and Schueler *et al* (2005)

References

Hugg T T et al 2020 The effect of sampling height on grass pollen concentrations in different urban environments in the Helsinki Metropolitan Area, Finland *PLoS One* **15** 1–12

Rojo J et al 2019 Near-ground effect of height on pollen exposure *Environ. Res.* **174** 160–9 Online: https://doi.org/10.1016/j.envres.2019.04.027

Schueler S et al 2005 Viability and sunlight sensitivity of oak pollen and its implications for pollen-mediated gene flow *Trees - Struct. Funct.* **19** 154–61

Schueler S and Schlünzen K H 2006 Modeling of oak pollen dispersal on the landscape level with a mesoscale atmospheric model *Environ. Model. Assess.* **11** 179–94