Reply to referee #2

I thank Andrew Barrett for his comments which have improved the manuscript. New manuscript text is italicized in the replies.

Comment 1: As a potential future user of this tool, I would appreciate a greater and clearer introduction of how the tool works. The language is still quite technical in places and certain terms (e.g. “objects”, “libraries” and “decorators”) are likely not well understood by future readers. While I acknowledge that there is extra information in the documentation of the tool, I suggest adding some more description to the paper. In particular I would like to see some more advice regarding the interfacing of the fortran code to the python tool. The current description sounds rather ad-hoc and I’m not clear how I would replicate this method.

Response: I modified Sec. 2.1 to be more didactic:

> *Two kinds of objects exist: those which represent a parametrization, and those representing the comparison. A standard object (an abstract class) is provided in order to define a parametrization (the PPPY box in Fig. 4). This abstract class already contains everything needed to perform the time advance and the saving of results but must be complemented (by inheritance) to incorporate the actual call to the different parametrization codes (Param1 and Param2 boxes of the figure). Finally, each parametrization can be used with different configurations. To achieve this, different instances (Param1.1, Param2.1 and Param2.2 boxes) are created, one for each of the configurations (e.g. time-step length, options specific to the parametrization).*
>
> *For the comparison, the provided class (PPPYComp in the figure) can be used directly or can be complemented (by inheritance, UserComp box in the figure) to add new diagnostics (e.g. new plot kind, computation of a derived variable to plot). An instance of the class is created for each comparison to perform (Comp box). A comparison is defined by the list of the parametrizations to use, the simulation length and the initial state. This comparison instance drives the parametrization instances to carry out the simulations and to plot the result.*

Moreover, an appendix is added to describe with more details the test example which is provided with PPPY. I think it is a good entry point to understand how the tool works before going through the other examples which are more interesting but also more complicated.

Comment 2: There is no description in the paper as to how this tool was used to identify and fix the causes of the time step dependence. The findings themselves are listed on page 9. Some additional description of how this tool enabled these model parameterization errors to be found and fixed is needed. This is the main benefit of this new tool, so it would be useful to see how it should be used.

Response: Please see comment #7

Comment 3: This paper is not the place to discuss in detail the merits of different (and in this case rather simple) sedimentation schemes, therefore I suggest shortening this section to allow for the above expansions.

Response: I reduced the sedimentation section.

Comment 4: The figures in this paper (which, I think, are produced by the PPPY tool itself) are missing units on all axes. Additionally, Figure 3 does not make clear which parameterization

scheme is shown in which panel. There two failings need to be fixed. Ideally, not just in the paper, but also in the code of the tool itself (for future users benefit).

Response: The unit on the time and altitude axis were indeed missing, I added them. And, I moved the unit of the plotted variable from the title to the y-axis (the PPPY user can already define the title and the x/y labels). In addition, I changed the title of the different panels of Figure 3.

Comment 5: There seems to be inconsistent model forcing used for the different schemes (or inconsistent physics regarding condensation in the different schemes). Comparing figures 1, 2 and 3 – the water vapor content (black line) for the shortest time step (1s) converges on a value of around 8 g/kg after 180 seconds, whereas for the other schemes in figure 2 and 3, the water vapor converges on 6 g/kg. What is the reason for this difference? Can it be corrected? I understand that the microphysics schemes will give different results for the hydrometeor concentrations, but the simple balance between temperature, water vapor and condensed water should be more or less the same for all schemes. At the moment it is difficult to compare results between the different figures/ microphysics schemes.

Response: In the ICE scheme, condensation is apart from the other microphysical processes. I first checked the time-step dependency in the condensation part. If condensation is activated (with one call to the subroutine by time step), it tends to hide somewhat the time-step dependency by making the different simulations to converge towards an equilibrium point. In particular with the ICE scheme with the same setup as used in the manuscript, the adjustment would suppress the cloud ice, and thus the time-step dependency of all the processes involving this specie will not be seen. Several setups would then be necessary to explore all the microphysical processes. But, indeed, with the adjustment, the vapour content of ICE with the same setup converges towards 6g/kg.

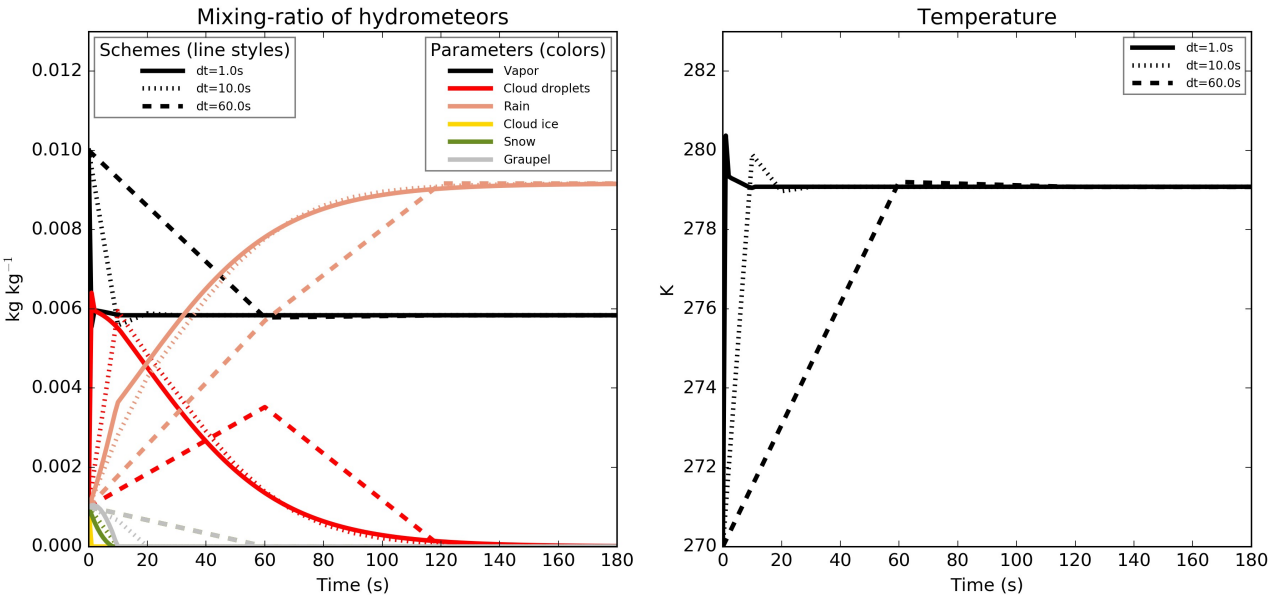

Illustration 1: Same as Fig. 1 of the manuscript but with the adjustment activated

For the LIMA scheme, adjustment is active, it would be easy to deactivate it. But, for the WRF schemes, the condensation process is embedded inside the microphysical parametrizations.

For the current study, I must deactivate the adjustment in ICE to fully explore all the microphysical processes. Because the goal is not to compare the results between schemes but to compare, for each scheme, the simulations for different time steps, I think it is not very important if the setup and/or the active processes differ. In a next study I hope to be able to compare the different schemes; in this next study it will be necessary to pay attention to the setup and to the active processes.

I made several modifications in the text to exhibit more this difference between the ICE scheme and the others:

- *(excluding the saturation adjustment and the sedimentation)*
- *In the simulations performed with this scheme (Fig. 2), the setup is the same as for the ICE scheme but the saturation adjustment is active.*
- *The WRF simulations are performed using the saturation adjustment included inside each scheme.*

Comment 6: page 5, figure 3. It is impossible to tell from figure 3 and the caption, which of the subpanels relate to which microphysics scheme. This should at minimum be added to the caption, and preferable to the figure panels too. The label "Schemes (line styles)" in the top left of each figure could easily be replaced by (e.g.) the scheme name/abbreviation for each panel. The comparison of different microphysics schemes in figure 3 is initiated with a rather unrealistic setup (approximate relative humidity is 165%). In full (3D/4D) model simulations, such supersaturation would never occur, and the microphysics schemes should not be expected to treat such situations fully realistically. Nevertheless, I find the differences between the schemes very interesting – I would be particularly interested to see how these same schemes performed in more realistic setups (e.g. with significantly reduced supersaturation at t=0 and/or when a constant cooling rate is applied)

Response: See comment #4 for the labels.

The setup used in the paper was chosen to allow for a maximum of microphysical processes to be active during a single simulation even if it is not fully realistic. Changing the setup by reducing the vapour content at t=0 (to use 4g/kg instead of 10g/kg) reduces the number of existing species for the ICE scheme (and hence the number of active processes) but does not suppress the time-step dependency. I think the figures with this new setup are less illustrative, I prefer keeping the old ones in the manuscript but you will find the new ones below:

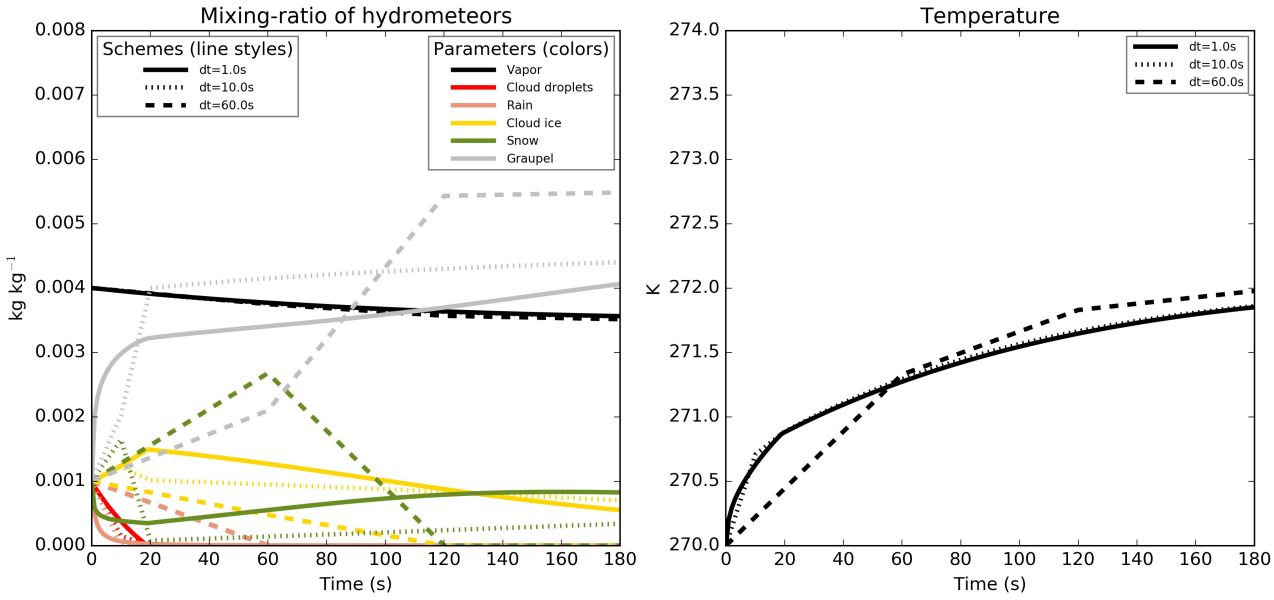

Illustration 2: Same as Fig. 1 of the manuscript but with a drier setup

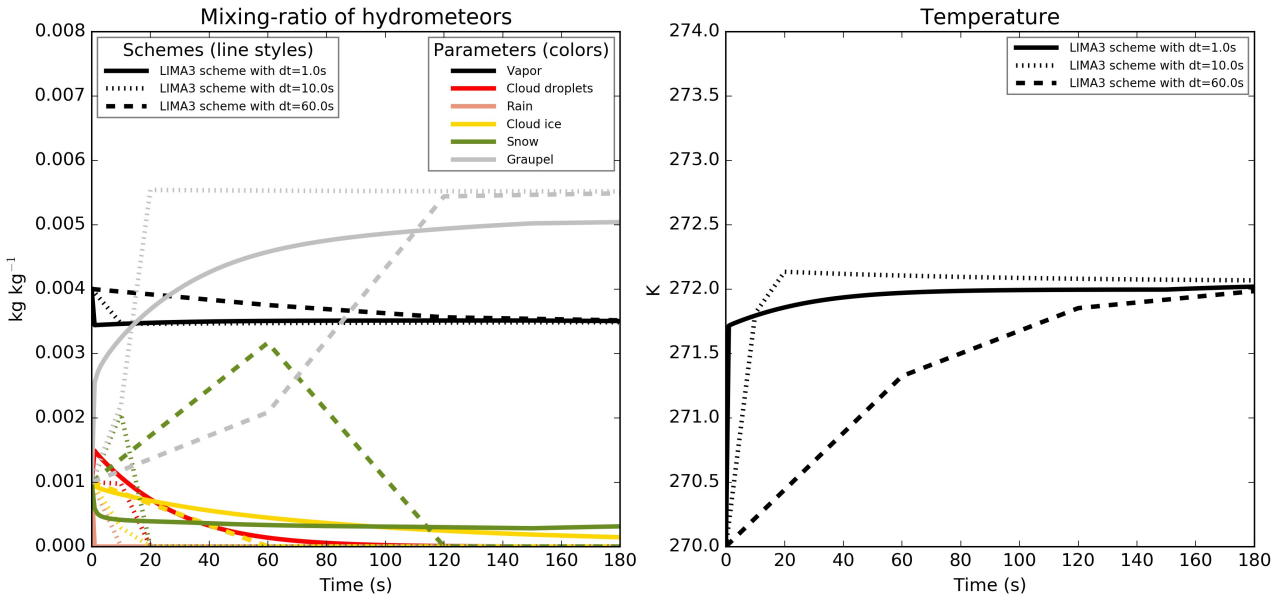

Illustration 3: Same as Fig. 2 of the manuscript but with a drier setup

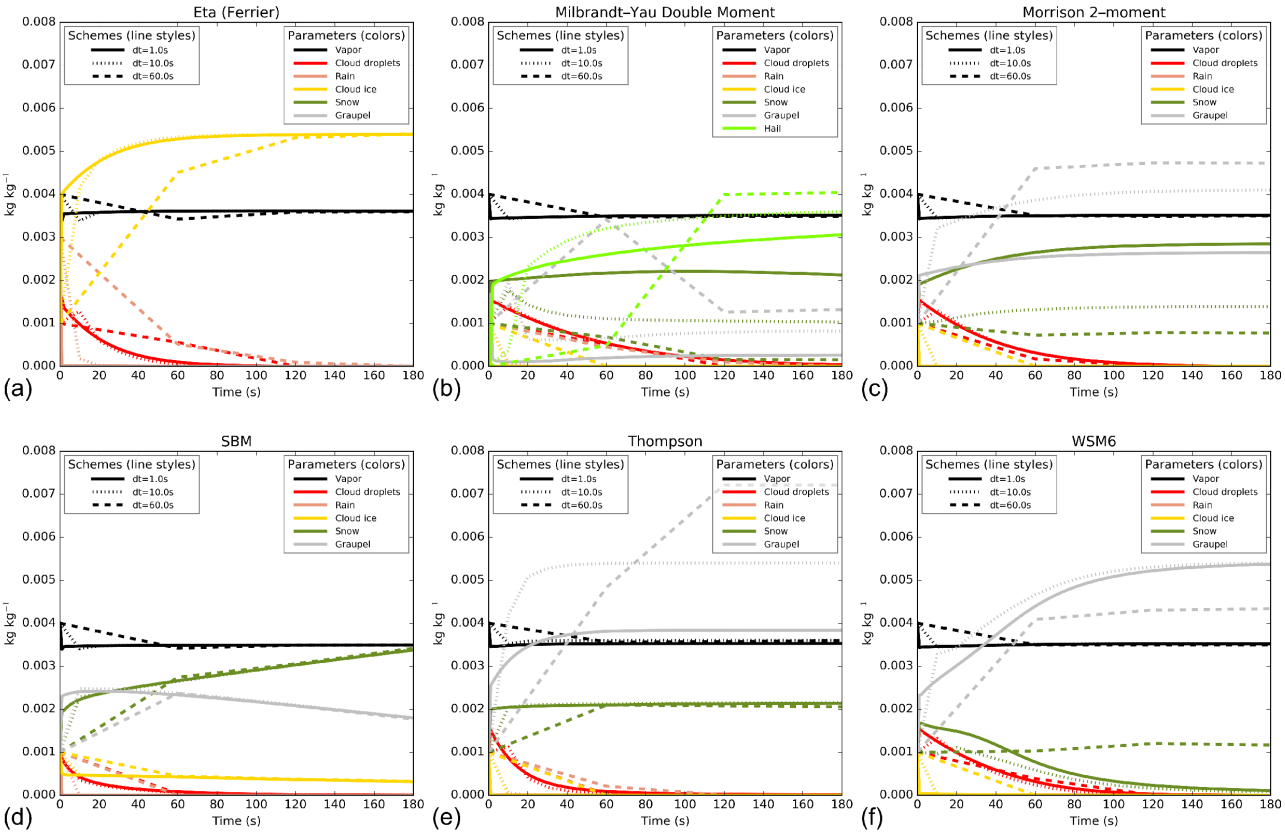

Illustration 4: Same as Fig. 3 of the manuscript but with a drier setup

I added a sentence in the manuscript to inform about the unrealistic setup:

> *The setup is not fully realistic (with an important supersaturation) but allows simulations to involve all the species and, hence, virtually all the microphysical processes. It was checked (not shown) that the time-step dependency still exists when using more realistic initial values.*

Comment 7: page 8, bullet points lines 9-28. These are all very interesting findings, but how did the PPPY tool help you to discover these factors as being important. It would be good to show the benefits of the tool you developed in achieving these findings. Was it simply a trial-and-error

process, or is there some aspect of PPPY that enables these errors to be determined more quickly?

Response: Unfortunately, I didn't find a better way than using a trial and error process (by enabling or not the different microphysical processes).

> *The simulations have been carried out several times activating and deactivating the different microphysical processes. To do this, the processes have been called individually by the PPPY software (when they are written in separate subroutines) or activated through switches or, at worst, (un-)commented in the source code. This trial-and-error process makes it possible to identify the processes that led to the oscillations and to the time-step dependency, and allowed the checking of each correction individually from the others.*

Comment 8: page 8, line 27-29 please clarify what you mean by "the conversion rate of graupel into hail is now computed from the wet growth rate of graupel and not from the total content of graupel". How large is this difference and why does it make such a difference?

Response: In the scheme the graupel is produced by the snow collecting liquid water, then the graupel growths by collecting other species (and with vapour deposition). If the graupel collects liquid species (rain or cloud), then there are two possibilities: the graupel is able to freeze the liquid content collected (this is called the dry growth mode) or not and a thin liquid film appears at the surface graupel (this is called the wet growth mode). The choice is made with the help of a heat budget.

> *the graupel growths mainly by collecting other species. When this collection implies liquid species (rain and/or cloud), there are two possibilities (called growth mode) depending on a heat balance: the graupel is able to freeze the entire collected liquid collected (dry mode) or a thin liquid film appears at the graupel surface (wet mode). In the original version of the ICE scheme there was confusion between the maximum content of liquid water than can be frozen (which must be used in the heat balance) and the content of liquid and ice water that can be collected in wet growth mode (which must be used to compute the graupel tendency). The correction made the mode choice more continuous. And because, ultimately, the graupel growth mode has an impact on the collection efficiency of icy species (snow and cloud ice) with the graupel, this choice can lead to significant differences in the collection rates. Hence, the scheme including the correction is less time-step dependent;*

If hail is activated, the wet growth mode contributes to the formation of hail. A conversion rate from graupel to hail is computed based on the collection rates and on the heat budget used to choose the growth mode. Then there are two ways of using this conversion rate:

- the old one: the conversion rate is applied to the entire graupel content. A given percentage of the graupel is then transformed into hail. For a same simulation length, depending on the number of time steps used, this conversion is made a different number of times. To simplify the reasoning, if the conversion rate is constant and equal to 0.5, the percentage of graupel converted into hail after 10s is 50% with one time step of 10s and 75% with two time steps of 5s.
- the proposed one: the conversion rate is only applied to the tendency of the graupel. The graupel already present at the beginning of the time step remains graupel and only the collected mass (in wet mode) can be converted into hail. This modification reduces the amount of hail produced and suppresses the time-step dependency.

The proposed version must still be validated but it is better than the old one because the time-step dependency is suppressed and because the old version had a tendency to produce small amount of hail under nearly all precipitating cold clouds (and this is no longer the case with the new version).

> *several modifications have been carried out on the processes involving the hail category as a prognostic field: the processes dealing with hail are now completely symmetric with those dealing with the graupel category (to ensure consistency even if this did not produce time-*

step dependency). A conversion fraction is computed from the heat balance used to choose the graupel growth mode. In the original version of the scheme, this fraction was applied on the total content of graupel; this induced a conversion tendency directly linked to the number of times the rate is applied (hence to the time step for a given simulation length). On the contrary, in the new version, the conversion fraction is applied on the wet growth rate, this way, no time-step dependency is produced. This was the main reason for the time-step dependency on the hail category.

Comment 9: Figures 11, 12 & 13. please make clear that the x-axis is timestep length (dt) – it could also be interpreted as timestep number (i.e. as a time-height plot)

Response: Thank you for the suggestion, it is done.

Comment: Technical/language corrections:

- page 2, line 9-10. please provide mode details about what differences were seen in the Meso-NH model when the time step was changed?

- page 3, line 9. what is a “tool package”. Where can the reader find it?

- page 3, line 13. “makes possible” → “makes it possible”

- page 4, line 1 correct to “consists of a python package written to ...”

- page 4, line 5 delete “which is the required”

- page 6, line 3 “needed to use the parameterization” → “needed by the parameterization”

- page 6, line 6 “that not exist” → “that do not exist”

-page 7, line 4 “to plug other” → “to plug in other”

- page 7, line 5 “one have to define” → “one has to define”

- page 8, line 2 “dependency on the simulation” → “dependency of the simulation”

- page 8, line 11 please quantify what you mean by “small time steps”- page 8, line 18 it is not clear to me what you mean by “graupel growth mode”, please give more details

- page 8, line 23 “has melt into rain” → “has melted into rain”

- page 8, line 25 “insure” → “ensure”

-page 9, line 10 & 11. please use standard scientific notation (e.g. 1.0x10 -5)

- page 9, line 11. please quantify the “substantial additional cost”

- page 9, line 16 “unique” → “single”

- repeated grammatical error (e.g. page 12, line 9; page 12, line 31; page 13, line 1, page 13, line 7): “Longer is the time step, more this part is important” → “The longer the time step is, the more important this part is”

- page 12, line 31. What is the “it” in “it reaches around 11%”?

- page 13, line 5 “mean content is weaker” → “mean content is less”

- page 13, line 20 “by consequences” → “as a consequence”

- page 13, line 24 “this induces” → “this means”

- page 13, line 25 “larger to one” → “larger than one”

- page 14, line 3 “None of both schemes” → “Neither of these two schemes”

- page 14, line 9 “weaker” → “less”

- page 14, line 28 “whatever is the time step” → “whatever the time step is”

- page 14, line 32 “hypothesis done” → “hypothesis”

-page 15, line 2 “certainly reduce” → “certainly reduced”

- page 15, line 10 “This scheme allows to make fall the bigger drops quicker” → “This scheme allows the bigger drops to fall more quickly”

Response: Thank you for the numerous corrections you have suggested. In addition, the paper have been reviewed by a native writer of English. Some of your remarks need a reply, they are below:

- page 2, line 9-10. please provide mode details about what differences were seen in the Meso-NH model when the time step was changed?
    - The test simulations have been done several years ago and are no more available. A rerun of those simulations would be necessary to give extended details on the differences.
- page 3, line 9. what is a “tool package”. Where can the reader find it?
    - *A documentation is provided with the software (see the code availability section).*
- page 8, line 11 please quantify what you mean by “small time steps”- page 8, line 18 it is not clear to me what you mean by “graupel growth mode”, please give more details
    - small time steps: replaced by *simulations using the small time steps (shown in Fig. 7)*
    - graupel mode: please see comment #8
- page 9, line 10 & 11. please use standard scientific notation (e.g. 1.0x10 -5)
    - It was an error, the values are 0.01 and 0.05
- page 9, line 11. please quantify the “substantial additional cost”
    - *The iterations needed for the 0.05g kg^{-1} threshold induce an cost increase of about 5%, and for the 0.01g kg^{-1} threshold an additional 20% can be expected (this last figure is an estimation because no sensitivity test have been performed on the whole domain).*
- page 12, line 31. What is the “it” in “it reaches around 11%”?
    - sentence reformulated: *For the 60s time step, 11% of the total water reached the ground.*