

Dear Referee:

Thank you very much for the careful review, which is very constructive. We have answered your comments, questions etc. point by point in below.

Best regards,

Clemens Wastl, and co-authors, 26.11.2018

Responses to RC1:

General Comments:

Does the abstract provide a concise and complete summary? No, the line 13 phrase 'lead to statistically significant improvements...compared to the original SPPT' is not supported by the figures that state statistical significance with respect to an ensemble with no stochastic physics. The abstract or the figures must be corrected in order to be consistent. It is acceptable to mention results that are not statistically significant, but they must be stated as such.

The sentence in the abstract has been modified accordingly. "... compared to a reference run without stochastic physics."

Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? Yes, figures 3-7 must be improved as requested by the other referee.

Figures 3-7 (including the captions) have been adapted accordingly.

Specific Comments:

page 2, line 10: clarify '(e.g. surface fluxes if surface tendencies are not perturbed)' because the SPPT concept could be applied to surface tendencies, although this has not yet been done (for technical reasons).

The phrase "... if surface tendencies are not perturbed ..." has been added to this sentence.

page 2, line 14: 'destroys the physical consistency' - what do you mean by consistency? Please clarify or delete. One could simply state that tapering means that model errors near the top and bottom are not represented.

We replaced physical consistency with "... physical consistent representation of model uncertainty in the vertical ..."

page 2 line 28: 'no interaction of the uncertainties between the schemes is considered'. To be fair, it should be more precisely stated that no interaction _inside the timestep_ is considered. As soon as you perturb the model state, all physics parametrisations will start acting differently at the next timestep. (intuitively the specific originality of iSPPT is that it perturbs physical interactions that occur at timescales shorter than the timestep)

Thank you very much for this comment. You are right, pSPPT and ipSPPT enable an interaction of uncertainties between the different parametrisations within one timestep, whereas in SPPT the model acts on the perturbations in the next timestep.

The phrase "... inside a timestep ..." has been added.

Technical corrections:

page 3 line 6: typo 'Meteorlogie'

Corrected.

page 3 line 10: add a mention to HIRLAM, who are actively developing AROME.

We added the phrase "... HIRLAM (High Resolution Limited Area Model, Bengtsson et al., 2017) ..." and the corresponding publication to the reference list.

page 3 line 12: Bénard (with an accent) (I seem to recall Bubnova has one, too)

Changed, also in the reference list.

page 3 line 24: you mean 'ensemble size', not 'ensemble spread'. Spread and ensemble size are different things, unless you use ensemble size to do something else that produces spread, in which case that 'something else' needs to be stated.

Yes you are right, we are meaning ensemble size.

Phrase changed to "... good compromise between ensemble size and computational costs."

page 4 line 14: append 'at each gridpoint'

Done.

line 15: correct 'defined by a tapering function(see below)' because you have not defined it yet.

Done.

line 16: delete the text between parentheses, because it is a repeat of line 7-8 (or the other way around if you prefer).

Done.

line 18: delete either 'the way' or 'how'.

Done.

page 5 line 5: move the sentence about tuning to section 2.3, because it s part of the experimental setup.

Done.

page 5 line 15: please explain in what sense SPPT was 'unsatisfactory' on the Austrian domain

The phrase "... which have produced unsatisfactory results within the Austrian domain ..." has been removed.

page 5 line 19: typo 'diretlcy'

Corrected.

page 7 line 5: typo 'variablese'

Corrected.

page 8 line 11: typo 'adopted' -> 'adapted' or 'adopted in'

Corrected.

page 8 line 16: rewrite 'supersaturation is translated in a slightly positive temperature increase' (also, 'positive increase' is a tautology)

The word "positive" has been removed.

page 9 line 5: the phrase 'model error is not noticeably influenced' needs to be corrected because (1) model error and forecast error are different things, and (2) error realizations can change even if their statistics do not. You probably mean 'ensemble average error (and/or average member forecast error) is not noticeably changed', since you seem to argue that the reliability of spread is improved by stochastic physics (which would be a valid statement).

Yes you are right, we are talking about ensemble average error.

Changed to "... ensemble average error ...".

page 9 line 28: insert 'score _differences_ ... in Fig 5 are much smaller'

Done.

page 10 line 18: again, replace 'model error' by 'ensemble average error' or 'average forecast error'.

Done.

page 10 line 19-20 replace 'negative CRPS difference' by 'a reduction of CRPS'

Done.

page 10 line 21: replace 'partially' by 'to a lesser extent'

Done.

page 11 line 8: replace 'at convective days' by 'on convective days'

Done.

page 11 line 19 insert 'tendencies _of_ U, V, T and Q'

Done.

page 11 line 20 correct 'to _a_ control ensemble'

Done.

page 11 line 24 'reveals a significant increase': a claim of statistical significance is a serious one, so it must be precisely expressed: with respect to what is the increase statistically significant ? SPPT or no SPPT ?

Thank you for this comment. The ensemble spread is statistically significantly increased compared to the reference experiment (Fig.6d).

We added "... statistically significant increased ensemble spread compared to the reference experiment."

page 12 line 6: insert 'analysis of a _set of_ convective events'

Done.

page 12 line 11: use a more precise term than 'critical', e.g. 'non-consistent', 'physically unsatisfactory', etc.

Thank you. We changed "critical" to "physically unsatisfactory".

page 12 line 12 replace 'switching off tapering' by 'tapering switched off'

Done.

page 12 line 13 grammar '_these_' perturbed fields'

Corrected.

page 12 line 14: 'could be switched off' is ambiguous', either write 'it is likely that it could be switched off' or 'we have switched it off'

The sentence has been changed to "...the tapering function has been switched off...".

page 12 line 34: 'is considered critical' is unclear, do you mean 'unsatisfactory' ?

Yes we are meaning unsatisfactory.

We changed "critical" to "unsatisfactory".

page 13 line 2: replace 'provides' by 'can provide' because microphysics do not always increase temperature.

Changed.

page 13 line 4: you do not really know that 'it cannot be assumed that T and Q have exactly the same error characteristic'. Better write 'it seems wrong to assume that T and Q have exactly the same model error characteristics'.

Changed.

References: please check the accents for all authors (e.g. Vié etc) a valid web address must be supplied for gray literature references (namely, Palmer et al 2009 and Szucs 2016).

Done.

Figures: besides the already mentioned problems with the labelling, the colours for pSPPT and ipSPPT in Fig 3-7 should be made easier to distinguish, perhaps by using a lighter blue for pSPPT. They are impossible to tell apart on two of my screens (not to mention their legibility for most colour-blind readers).

Thank you for this comment. We changed the color for pSPPT to a lighter blue.