Replies to Referee Comments on "Evolution of squall line variability and error growth in an ensemble of LES " [1]

Edward Groot, Holger Tost

September 20, 2022

1 General

As authors we would like to thank both referees very much for their thoughtful reading of (and extensive comments on) the manuscript.

2 Referee 1: reply to RC1 [2]

2.1 High level comments

We thank referee 1 for filtering out and formulating his pointwise summary, with which we also think most of the important outcomes have been covered [2].

RC1: "(...) but rather a sharp eye and crisp tongue about what is important. The slow resurfacing at the end into Synthesis and then Conclusion sections (also rather long) does help to pull these key points and highlights out somewhat, but those key points could be even more polished into the Abstract for instance. (...) But those wrapup sections (4 and 5) come after a long slog of sometimes unclear prose in the late-middle (the very long sections 3.3 and 3.4)." and *RC1:"Might the paper or at least sections 3.3 and 3.4 be cut by aspirationally 50% with no loss (and a gain of clarity) on the reader's part? Long stretches of text appear to describe figures not shown, without stating (not shown). "*

In general we do agree that the manuscript is obviously very lengthy at the moment. Therefore we will move considerable parts of Section 3.3 that are less crucial for the story line of the manuscript to the supplementary material, so that highly interested readers can access it easily, but that it does not add extra load to the main text for any other readers. Some parts will even be removed completely in the revised version.

RC1: "Lines 19-20: The meaning of this result is that the perturbations chosen are in a non-essential field, but that even those differences grow (or explode). What does the word "intrinsic limit", lifted from some over-realm of philosophy it seems, really add to this idea?" and "480-onward: "error" —> "difference" "

Regarding Section 3.4 [2]: referee 1 sees the importance of investigation of ensemble spread in LES simulations and our focus on convective dynamics, which is indeed the main point. However, these comments do suggest that its relation to "errors" in the representation of the atmosphere within a model, as well as theoretical limits of predictability, are seen as comparatively unimportant. Maybe we overinterpret referee 1's criticism by reading it in RC1 this way. From our point of view there is no need to judge on which perspectives are important to investigate or not - however, as authors we do think it is necessary to look at atmospheric modelling and predictability from this perspective, as [3, 4], [5, 6, 7] and other work by [8, 9, 10] has shown; this implies that Section 3.4 will not be as strongly modified as Section 3.3: even if "error growth" and "intrinsic limit of predictability" might never be important in studies of LES ensembles of convective systems, we choose to represent the error growth point of view in this work in line with works by colleagues in the field [3, 4, 5, 6, 7, 8, 9, 10] .
Moreover, terms as "intrinsic (limit of) predictability" and "errors" have been used very regularly in a set of recent predictability studies in the footsteps of Lorenz' work [e.g. [11]] and applied to various synoptic and convective scale case studies and more climatological studies [3, 8, 9, 6, 10, 7] (all of them subscales of the atmospheric system that Lorenz' work described). This was done in order to distinguish between practical

predictability in state-of-the-art NWP and explosive/exponential error growth with its implications for predictability from even smaller scales down to the smallest possible scales and/or larger scale differences in state of (much) smaller amplitude: the practical predictability representing current state-of-the-art (operational) global NWP and intrinsic predictability stemming from the smallest (scale/amplitude) errors. This means that it is not [2] *""intrinsic limit", lifted from some over-realm of philosophy"* as formulated by Referee 1 in RC1, but rather directly addressed in (recent) meteorological studies. The cited papers (to the authors' opinions) depict why and how of given terms in the field of atmospheric modelling and predictability (of which not all atmospheric modellers will understandably be aware).

RC1: "From line 500 or so: the paper is getting rather long and verbose... can it be streamlined? If all the key results have been shown, why not gather them crisply and close?" and *RC1: "but those key points could be even more polished into the Abstract for instance."* and *RC1: "how many words could be trimmed or eliminated without loss of meaning?"*

On top of that the authors will pay attention to the writing style throughout the manuscript - reconsiderations in the discussion and conclusion sections will be done (including urgency of some paragraphs and sentences), which will be done in convolution with comments from RC2 [12].

RC1: "An unclear overanalysis in confusing statistical terms (terms like "source" and "error", and "autocorrelation" for intra-ensemble rather than temporal-lag correlations) were unhelpful or confusing, and it all gave few clear insights that aren't in the bullets above and in section 3.2."

Lastly, the use of the other "confusing terms" [2] is reconsidered (and mostly adjusted) for increased manuscript clarity. We hope to meet the expectations of referee 1 with the way we compact the (main) text and at the same time we hope to warranty completeness of the manuscript to any reader.

2.2 Local comments

We thank the reviewer for the local comments and will address them one by one in detail when updating the manuscript. Some specific comments (questions) will be adressed below.

- *RC1: "L158: "interface height" — this is just a reference value in an analytic formula, which translates into shear strength on the 100m grid, right? This description was quite confusing. "*
  We refer to Section 3.
- *RC1: "L253: does your contouring routine treat \sloped features different from / sloped features?"*
  There is no difference.
- *RC1: "L399: "trough" and "crest" — what do these mean? Is this the u field, does it even show vertical displacements at all?"*
  Yes, we see undulations, which coincide with perturbations in the u and w fields that show upward (downward) displacements and convergence/divergence patterns suggesting a through or crest.
- *RC1: "L433: "compensating" — what does this word mean? It implies a big back-story in the authors' minds about how things are related and constrained, makes me nervous."*
  There is not a really large "back-story", which means that we will choose other words.
- *RC1: "L442-444: huh, how does removing all buoyant gridcells remove "gravity wave contributions from saturated parcels" ?"*
  We think that the updated version will clarify this issue.
- *RC1: "L408: "circulation" — what does it mean? We are looking at complicated multi-lobed structure of the u field."*
  The convective overturn with main updraft and downdraft is meant with that. It includes a deeper overturn, but also shallower overturn and some entrainment (detrainment). In the revised manuscript circulation is sometimes replaced by (relative) u, flow anomalies/perturbation(s), etc. Essentially these refer to the structure found in Figure 8a/8b [1].

The comments that are not mentioned in the list have been addressed in 2.1 or are addressed below, in Section 3. Or in a few cases the figures or text in the manuscript will be updated or reconsidered for the revised version accordingly.

3 Referee 2: reply to RC2 [12]

3.1 Major comments

RC2: "vertical grid spacing: while the authors invest a lot of computational resources into running an ensemble at a high horizontal resolution, the vertical grid spacing is relatively poor. The equidistant spacing of 100 m in the vertical is in my view inadequate to resolve the cold-pool dynamics and maybe also the melting layer properly. Especially as the cold pool plays a vital role for the further spread between ensemble members, a fine grid seems to be critical to resolve differences in the evolution of the cold pool, its interaction with the flow and the feedback onto the squall line dynamics. I suggest to rerun the control simulation with a vertically stretched grid and to document the differences in cold-pool dynamics with the equidistant grid."

Regarding the vertical resolution, the selected choice of resolution has not directly been clarified and motivated within the manuscript.
We agree, that 100m vertical resolution may be inadequate to properly resolve the melting layer and also that cold pool dynamics might be affected by processes on smaller scales, such that they cannot be properly resolved with a layer thickness of 100m. However, the focus of this study was not on the cold pool dynamics, but the evolution of the ensemble spread of an idealised squall line with relatively small disturbances. Therefore, the aim is to identify aspects of convective systems and their dynamics that lead to variability and ensemble spread and thus how these aspects affect predictability, but not how smaller scale processes throughout the cascade are perfectly represented or not - an arbitrary length scale is setting the representation threshold for analysis, which in our case is somewhere near 1-2 km: cold pools can in principle just be represented.
Although we do see the point of referee 2 about vertical resolution, the concerns are from our point of view much lower. As the cold pool depth initiates at about 2.500 m and evolves to depths of 1-2 km (# grid cells 10-20), this appears to be close to but not yet at the critical low end of (cold pool) representation (about 7, but definitely at least 5 grid cells for any oscillations [13]) to the authors' opinion. Furthermore, the 100 m vertical resolution is very reasonable for representing squall lines and cold pool triggered systems in our opinion, based on studies by [14, 15, 16]. For the processes in the melting layer though, we would agree that the simulations are more or less *at* the critical low end. Indeed, comparison of the ensemble to a simulation at finer vertical resolution at low levels could therefore be seen as desirable, but as a short remark or paragraph in the **(supplementary?)** material, which we will add.
That should be sufficient, as the representation of very subtle microphysical processes are not exactly within the scope of this study. Somewhat suboptimal, poorer representation of shallow layer microphysics and cold pool dynamics will possibly induce very similar biases in all ensemble members and we do not see any reason to assume that the ensemble spread is "blown-up" overly.
Having executed a simulation at 50m vertical resolution with improved cold pool representation, the reproduced copy of Figure 3 of [1] with the new simulation is very similar to Figure 3, with (in terms of details) intermediate behavior between the control simulation and ENS-03. Hence, it may be concluded that the cold pool propagation and dynamics has to fall somewhere in the middle of the ensemble envelope.
Of course the validity of the comment remains and we appreciate the referee for triggering our thoughts about this topic.

RC2: "Please provide more information about the employed tracers, for example as a further subsection in section 2. How are they transported by the flow, in which way is the coupling to microphysics (sedimentation) and turbulence realized, what is the treatment of the tracers in the surface layer? The last point my be especially important as the lowest atmospheric layer is very deep, and without a careful treatment tracers may be stuck in the surface layer."

Moving to the next major comment, the tracer is passive also with respect to clouds and precipitating particles. It is not sedimenting. The tendencies acting on the tracer are purely advective, plus the subgrid turbulence from the LES (will be mentioned in manuscript update). Some tracer mass can indeed get stuck in the boundary layer (probably meant with surface layer?), which was illustrated by the motion of 95-99% of the tracer mass when installing another tracer (not included in the manuscript). Once tracer mass escapes from the boundary layer or the layer below the inversion through updrafts, it can move away from that layer. This is what happened with the large majority of tracer mass initialised in front of the cold pool edge. These are the tracers included in the manuscript. Some of the mass of these two tracers could indeed return to the lower (near-surface/sub-inversion) layers after a fast, small scale circulation in the convective drafts. This has all been taken into account, but the tracers represented in the manuscript are those that undergo motion towards the convective region at $x = 0$ km (initially).

RC2: "The Weisman and Klemp (1982) sounding is established and popular. Yet, it has been criticized for being very unstable and favourable for convection. For the case at hand this means that in all ensemble members a squall line develops. A profile that is less favourable for convective development the ensemble spread may be much larger, as some members may not be able to produce a vivid squall line."

With regard to the comment about the Weisman and Klemp sounding: this is very true. The study we present assumes from a practical point of view that a line of strong updrafts initiates and this is certainly not always the case in an ensemble of simulations when squall lines might form, which has a large control on ensemble variability as well. Nevertheless, we decided to use this scenario, as the variability of interest is not whether a squall line develops, but how the spread in the potential squall line state evolves. Therefore, it represents a subsample space in potential squall line variability.

Of course, convective initiation by itself is a similarly interesting topic, but not in the scope of the presented study. From the other point of view (variability asssociated with the question "does deep convection initiate?"), earlier studies like [3] are offering a more insightful perspective and this study does not address that point of view. In other words, we argue that the preferred sounding depends on one's interests and thus point of view.
Since the case of no convective initiation would add a (0,0) point in many spaces of interest (i.e. no mass flux, precipitation and no flow anomalies as a result of deep convection), the extractable statistical signals are likely even highly influenced by cases of untriggered convection, burying the actual signal that is identified in this study. To the opinion the authors, looking at ensembles of convective systems in an integrated way that considers both perspectives however, could be a way to go for future theoretical studies, for which we have some ideas.
The initiation-conditional point of view will be pointed out more explicitly in the updated manuscript.

RC2: "The reference simulation appears at one end of the spectrum, while I would have expected it somewhere in the middle of the ensemble. Do you have an explanation for this behaviour?"

There is not supposed to be any preferred location with regards to the location of the reference simulation within the ensemble envelope. Initial condition perturbations are imposed on the thickness and top altitude of the shear layer, as described in the manuscript (and then translated to the model grid, as it will be addressed in more detail in the revised manuscript). Magnitudes of these perturbations vary from -5 to +1 or 2 % compared to the reference simulation, where the shear layer height in the reference run is 100%. The magnitude of the initial perturbations does not correlate with the perturbations at later times, as gravity waves are found to decorrelate those perturbation signals within the first 30 minutes.
When the convection is active and the cold pool accelerates, at about 30-40 minutes into the simulation, a perturbation structure that is maintained exists on a time scale of about an hour. After that, another episode of decorrelation starts.
The decorrelation phases destroy the structure in any initial condition perturbations and hence it is not known how any initial condition perturbation will evolve in this first stage of about 30 minutes (in our set-up). Predictions can only be built from the state after 30 minutes or so, towards states in the next hour. We tried to discuss this in the discussion section of the manuscript, but obviously, this section is going to be re-written in the updated manuscript to better elucidate this aspect.

3.2 Minor/local commments

For the updated manuscript, following is done in addition:

- *RC2: "subsection 2.1: please give more details about the formulation of microphysics, especially the treatment of the condensation process (via saturation adjustment ?) and the evaporation process, as they will be crucial for the development of up- and downdrafts and thereby ensemble spread."*
  More microphysical details, including an extended description, references and code version, will be added in an updated 2.1
- *RC2: "Please provide more information on the perturbation of the initial conditions, especially on the perturbation of z_i. In which way is the perturbation of z_i transferred to the atmospheric profile."*
  A short description of how perturbed interface heights are translated to winds at the model grid will be provided
- *RC2: "The top of the model domain is at 20 km, with a sponge extending down to 15 km. Taking a look at e.g. Figure 8 some of the convection seems to interact with the sponge already. Did you see*

any signals of interaction of the convection with the sponge?"
The convective cloud tops "live" at elevations of 10-13 km roughly. That means a very low fraction of convective clouds will exceed 13 km and the absolute limit that updrafts reach is at about 14 km. Similarly, the divergent upper tropospheric motion does stop at about 13 km too (Figure 8). There is a 1-2 km zone between the top of convective updrafts and the sponge layer. In the sponge layer itself only footprints of gravity wave motion are seen, as one would expect (sponge layer is there to damp these motions). Otherwise no clear dynamical effects are seen in this sponge layer. Therefore, there are no indications of (undesired) interactions of clouds and updrafts with the sponge layer.

- *RC2: "Figure 1: Please specify the computation of the parcel ascent. The red line seems to start at some elevated point. Please also remove the "(left)" statement."*
  Displayed is indeed the ascent of a parcel from the mixed layer at about 900 m altitude in the lower troposphere, as it will be described in the revised version.

- *RC2: "Section 3.1: please provide some detail about the computation of the radar reflectivity"*
  The computation is entirely based on CM1 output, which uses the specific humidity (cloud contents) in the grid boxes and the computation will be shortly described in the updated manuscript.

- *RC2: "section 3.3.1: there is some directional shear in the simulations given by the increasing v velocity component. The averaging over the y direction ignores this directional shear. In which way did you account for this?"*
  This minor asymmetry has not directly been accounted for, which is because the structure of the squall line remains highly linear/2D. In combination with the along-line spatial average, this is sufficient to extract the signals that are presented in [1]. Small corrections as suggested by [12] would have tiny impact on top of the spatial averaging.
  The meridional wind has been triggered to ascertain the manifestation of some 3D turbulence, as it has been described in the manuscript [1].

- *RC2: "Section 3.3.2: please give more detail on the ensemble sensitivity analysis. The section is impossible to understand without first taking a look into the cited papers. The 4th and 5th paragraph of the subsection is hard to follow, there is no figure supporting the statement "during the first 15 minutes of simulation time ..." and "After 15-20 minutes" "*
  Based on referee 1 [2], some of this unclear material will be moved to the supplementary material. The corresponding figures can also be found in the current version of the supplementary material.
  And: additional content on the ensemble sensitivity analysis will be provided in an updated version of the manuscript.

- *RC2: "Section 3.3.3 downdraft selection: by selecting only grid points that contain hydrometeors, the downdrafts where all rain has been evaporated will be disregarded. A better choice could be to increase the magnitude and/or to check for hydrometeors above."*
  The authors agree with the referee [12] that detecting columns in which evaporation takes place might be a better choice for detecting downdrafts. However, we think that this choice is not important to our results. We are not aware of any study comparing various algorithms with small differences like this for downdraft detection in a high resolution simulation of deep convection. A thorough comparison of detection algorithms would be a study on its own. In such a study a detailed comparison would definitely appropriate, including a test and evaluation of other aspects of the detection algorithm. The question of whether this suggestion would improve the downdraft detections is therefore well beyond the scope of this study and only considered relevant by the authors in a study where downdrafts play a (much) larger role.

- *RC2: "Line 461-467: I cannot follow the argumentation here. Judging from Figure 10 the downdraft mass flux at a height of 1.5-2km seems to show the largest variability. Please be more specific or rephrase."*
  The section describing the low level mass flux variability in downdrafts within the ensemble (Figure 10) will be updated, reconsidering the need for clarity.

- *RC2: "Line 590-593: I disagree with this statement. Downdrafts will carve their space, irrespective if there is space available or not, thereby killing updrafts."*
  These sentences will also be reformulated into something less suggestive, or we may consider removing a small portion completely (see also RC1 [2]).

Furthermore, the authors will consider and address the technical points of referee 2 to further improve the quality of manuscript.
We highly appreciate the contributions by referee 2 to improve the quality of the manuscript.

References

[1] Edward Groot and Holger Tost. Evolution of squall line variability and error growth in an ensemble of les. *EGUsphere*, 2022:1–34, 2022.

[2] Referee comment 1 on "evolution of squall line variability and error growth in an ensemble of les", 2022.

[3] Christopher Melhauser and Fuqing Zhang. Practical and intrinsic predictability of severe and convective weather at the mesoscales. *Journal of the Atmospheric Sciences*, 69(11):3350–3371, 2012.

[4] Jonathan A. Weyn and Dale R. Durran. The dependence of the predictability of mesoscale convective systems on the horizontal scale and amplitude of initial errors in idealized simulations. *Journal of the Atmospheric Sciences*, 74(7):2191 – 2210, 2017.

[5] L. Bierdel, T. Selz, and G.C. Craig. Theoretical aspects of upscale error growth through the mesoscales: an analytical model. *Quarterly Journal of the Royal Meteorological Society*, 143(709):3048–3059, October 2017.

[6] Tobias Selz. Estimating the intrinsic limit of predictability using a stochastic convection scheme. *Journal of the Atmospheric Sciences*, 76(3):757–765, 2019.

[7] Tobias Selz, Michael Riemer, and George Craig. The transition from practical to intrinsic predictability of midlatitude weather. *Journal of the Atmospheric Sciences*, 2022.

[8] Fuqing Zhang. Dynamics and structure of mesoscale error covariance of a winter cyclone estimated through short-range ensemble forecasts. *Monthly Weather Review*, Oct 2005.

[9] Fuqing Zhang, Naifang Bei, Richard Rotunno, Chris Snyder, and Craig C. Epifanio. Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *Journal of the Atmospheric Sciences*, 64(10):3579–3594, October 2007.

[10] Fuqing Zhang, Y. Qiang Sun, Linus Magnusson, Roberto Buizza, Shian-Jiann Lin, Jan-Huey Chen, and Kerry Emanuel. What is the predictability limit of midlatitude weather? *Journal of the Atmospheric Sciences*, 76(4):1077 – 1091, 2019.

[11] Edward N. Lorenz. The predictability of a flow which possesses many scales of motion. *Tellus*, 21(3):289–307, 1969.

[12] Referee comment 2 on "evolution of squall line variability and error growth in an ensemble of les", 2022.

[13] William C. Skamarock. Evaluating mesoscale nwp models using kinetic energy spectra. *Monthly Weather Review*, 132(12):3019–3032, 2004.

[14] George H. Bryan, John C. Wyngaard, and Michael J. Fritsch. Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, 131:2394–2416, 2003.

[15] George H. Bryan and Hugh Morrison. Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Monthly Weather Review*, 140(1):202 – 225, 2012.

[16] Leah D. Grant and Susan C. van den Heever. Cold pool dissipation. *Journal of Geophysical Research: Atmospheres*, 121(3):1138–1155, 2016.