Replies to all Referee Comments on "Divergent convective outflow in ICON deep convection-permitting and parameterised deep convection simulations. " [1]

Edward Groot, Patrick Kuntze, Annette Miltenberger, Holger Tost

February 6, 2024

1 General

We thank both referees for reading our manuscript and formulating their summary and feedback [2, 3], which we welcome to make further improvements to the manuscript. We think that the reviewers have provided very useful input to do so, which brings us to some purely textual revisions and a figure correction to accomplish an improved manuscript and address the concerns of the reviewers.

2 Referee 1: reply to RC1 [2]

1. *'"The authors have made a substantial effort to significantly improve their manuscript on dimensionality, ensemble and convective momentum transport aspects and it is now acceptable, but I suggest to address the remaining minor points below (including shortening section 7)"'*

   We thank the reviewer for this feedback and are very glad that the content of the work has been much clarified after the revisions.

2. *'"l59 typo "In (?) explicit ";*
   l82 "is inversely proportional to their vertical wavelength" as $m^s = N^2/C^2$, speed C is inversely proportional to vertical wavenumber or proportional to vertical wavelength, please double check and correct;
   l247 "to quantity geometry" you mean "quantify the"?
   l397 typo "symbolsIn"'

   The updated manuscript will be corrected accordingly.

3. *'"I found section 7 and Conclusions very long and sometimes repetitive, one could remove 7.3.2 and shorten 7.2.1 by half"'*

   We will have another critical look at Sections 7.3.2 and 7.2.1. We would also just like to remind the reviewers of the earlier round of reviews, in which Section 7 was considered critical and potentially highly relevant to bring up novelties in this study [2, 4], but at the same time a section to extensively work on. Therefore, we will also try to beware when shortening these parts of Section 7, but we agree that 7.2.1 is particularly lengthy and shorten this Section. Section 7.3.2 can probably be merged with the previously existing 7.3.3 in the revised manuscript. We would also like to thank the reviewer for raising this point.

3 Referee 3: reply to RC3 [3]

- *'"I find this article a very interesting read and I hope that it is published. The relation between net latent heating (quantified as the surface precipitation rate) and divergent outflow in the upper troposphere is investigated, using ICON simulations of a real case of severe thunderstorms over Southern Germany. This builds on the author's idealised LES study already published. Consistent with the earlier LES work, the present study finds that upper-level divergent outflow increases with latent heating, but* *increases more slowly as the organisation / clustering of the convective updrafts increases. This effect is found in 1km resolution simulations, but is absent from simulations at 13km resolution. It is suggested that resolving small-scale gravity-waves is key to representing the relative slowing of upper-level divergence with increasing convective organisation. The relation between upper-tropospheric outflow and convective momentum transport is also investigated. "'*

We would like to thank the reviewer for this very encouraging view on this work. Furthermore, we are glad that the key message could be distilled from the work, which had turned out to be rather difficult from the initial submission [2, 4].

- *'"However I have a number of fundamental questions about the theory, methodology and how the results are interpreted. I hope that in answering these questions, the paper can be made clearer and stronger. In its current form, it is not clear to me that all of the stated conclusions are proven by the results shown.*
  So I would recommend significant revisions, if that is still possible at this stage? I also note that this paper was originally submitted 9 months ago, so has been held up for a long time already (I understand one of the previous reviewers is now unavailable, which is why I have been invited to review the article so late in the process). Many apologies if my queries introduce further delay"'

  We will address the concerns in the following point by point and, wherever we think it is needed, include the relevant considerations into the discussion of a revised version of the manuscript. We think that addressing the reviewer's questions will indeed make the work stronger. Furthermore, we are glad that the reviewer is aware of the specific setting in which a review of our work has been requested.

1. *'"Introduction: L20 and onwards. A key argument of this paper is that the amount of upper-level divergent outflow driven by a given amount of latent heating within convective clouds varies as a function of storm horizontal structure. However, surely the total (time-integrated) amount of ascent (and hence upper-level divergence) is constrained to be that required to restore the heated air to neutral buoyancy? This is the basis of "Weak Temperature Gradient" scaling argument, which is key to the dynamics of moist convection and its interaction with the stably-stratified free troposphere. Latent heating will temporarily make the cloudy updraft air buoyant relative to its surroundings; it then ascends relative to its stably-stratified surroundings until it reaches its neutral buoyancy level, and the amount of upper-level divergence forced by this must be constrained by mass-continuity. If the time-integrated upper-level divergence was less than this constraint, then the cloud column would remain buoyant after ascent had stopped. This should only be possible if the resulting warm anomaly is held in place by a rotationally-balanced flow, which will not occur at convective scales. I think the paper's findings can be reconciled with this argument because it actually considers the transient response of the upper-level divergence to the latent heating, not the total time-integrated response. We would expect the upper-level divergence to continue after the latent heating has stopped, until the warm anomaly is removed, but this final stage of the outflow is not considered in this study. So, the authors find significant variation in the transient rate at which the upper-level divergence spins-up in response to latent heating, as a function of storm structure. This is a very interesting result, especially the discrepency of this between different model resolutions. I would just ask the authors to clarify that their results apply for the transient response and not the time-integrated total response of upper-level divergence to latent heating. e.g. replace the often repeated term "amount of upper tropospheric divergence" with something like "instantaneous rate of upper tropospheric divergence""'*
   We agree that the distinction should be made between a time-integrated and transient response and the terminology is adapted accordingly for the revised manuscript and detailed in lines 74-81 of the revised manuscript. The time-integrated response in the setting of a realistic NWP simulation requires a different methodology and approach, because systems at time lags are indeed thought to produce time-integrated divergence over larger mesoscales with mixed transients, consisting of infragravity wave signals that are in different adjustment states regarding their transients to balanced flow [e.g. [5, 6]].
   We have tried to investigate if there was some lag effect in the response of transient divergence signals in [7] affecting the time-integrated response on long time scales, but have not been able to find substantial evidence for that. However, in these idealised two hour simulations, the Coriolis effect has been switched off, which clearly controls the time-integrated effect [7].
   Given the realistic and more complicated setup of the current study, the integrated effect is not what we aim to analyse here and should indeed be emphasised in this context (also by shortly addressing this point in the introduction and/or discussion).

Indeed, by focusing on the transient, localised and few-hour time scales, we distill responses to convection that are relevant to transient momentum and heating transfer to larger mesoscales (MCS scales and slightly beyond; as quantified by divergence) and their variation in representation.

2. *'"One factor not considered in this paper is the influence of the vertical structure of the latent heating profile on the mass divergence rate. If heating is concentrated in a layer of the troposphere with higher (lower) static stability, then less (more) ascent is required to restore the heated air to neutral buoyancy, and hence less (more) upper-level divergence will result. One possible hypothesis that might be worth checking is whether part of the variation in upper-level divergence per unit latent heating can be explained in this way by variation in the vertical structure of the heating profile, rather than storm horizontal morphology. It might be that the coarse simulations have deficient variability in the vertical structure, and hence give less variability in outflow strength"'*

   These hypotheses have not been ignored for our works (both [7] and this study; which will be explained in the following paragraphs). In addition, if we assume that the warm air mass is relatively homogeneous (to the east of the frontal zone, which is over the French-German border region) and that the convection takes place on the flanks of this warm air mass, there may not be sufficient variation in stability profiles among all the simulations to cover this aspect of vertical variability with our case study (studying the convective systems in the warm air mass over Germany, directly east of the frontal zone). It should not be an issue of the different ICON configurations, as the level of neutral buoyancy will typically locate at or very near to the tropopause, with nearly consistent and nearly similar levels of neutral buoyancy and maximum divergence between both configurations (e.g. Figure 6). However, in the below, the impressions based on the two studies are summarised.

   We have looked into this in our previous work, [7]. There, we induced variation in the heating profiles by manipulating the constant of latent heat release. We have also tried this procedure in ICON(-PAR simulations - with physics manipulation), but in both [7] and the current case, we could not find structural responses to differences in heating profiles other than those relating to variations in the level of neutral buoyancy. This response was clearly visible in the idealised work of [7] and less clear for certain systems in the current study. Some systems show a clear response of differences in outflow levels and others show a response that is statistically negligible. Also, in both [7] and the current work, the magnitude of (instantaneous) divergence is suggested to not be structurally affected in all those simulations (at least not statistically). It should be noted that the idealised setting allows one to distill the statistical signals in a more confident way than the configuration of the current study.

   As a result of the above findings, we think that the way it has been reported in [7] covers what we can confidently say about outflow level/thickness variability among the convective outflows.

3. *'"Presumably the gravity-wave dynamics thought to mediate this impact could be at least partially resolved at this resolution, and represented by the model's dynamical core rather than the convection parameterisation? "'*

   First, we need to disentangle the production of the (instantaneous) divergence: the convective parameterisation and/or grid-scale microphysics produce net heating as a consequence of net condensation and precipitation. As a result, the troposphere will warm. The dynamical core is informed by the accompanied heating. Therefore, a response is triggered in the form of the divergence at upper levels. Furthermore, convergence at low levels can be enhanced.

   '"L136: Why would it not be possible to represent the impact of storm morphology on upper-level divergence patterns in a simulation at 13km resolution with parameterised convection? (...) I similarly query the discussion at L439-440: "little or no information on geometry of the convective systems can be represented with a parameterisation"; surely much of the geometry (as quantified by the ellipse fitting in this study) is mesoscale, so occurs on the resolved-scale in the 13km grid PAR simulations. So the gravity-wave response to the storm structure could be captured. The results shown in figure 8b show that the divergence-precipitation relationship is not sensitive to the choice of parameterised versus explicit convection at 13km resolution, so the paper's argument that the overly-linear response in PAR is due to the convection parameterisation (e.g. L680, L710) is not supported. It might be that the PAR simulation simply fails to produce enough emergent mesoscale organisation on the grid (even though it could be resolved). However, figure 7(c,d) appears to show plenty of mesoscale structure in the divergence field for the PAR simulation."' and *'"In various places the paper claims that the dependence of outflow strength on convective organisation cannot be represented when using parameterised convection. But the results don't support this conclusion. The organisation quantified using the fitting of ellipses to the precipitation field is largely on scales that are resolvable in the parameterised convection simulation. Any resulting gravity-waves connected with the mesoscale heating structure would similarly be resolvable. And the results in figure 8b show that turning off the convection parameterisation (while keeping all else equal) made no difference to the divergence-precipitation relation"'*

We do agree that the statements in the paper could be interpreted as overstatements of the contrasts, which is indeed suggested by the data points in Figure 8 that compare the 13 km ICON with and without convection parameterisation. We have revised the implied causalities and methodological descriptions along this line, correcting from *"that convective organisation is not directly represented in simulation configurations with parameterised deep convection"* to ***"that convective organisation is weakly represented in simulation configurations with parameterised deep convection and weakly coupled to the engine of numerical models (the dynamical cores); we could say it is clearly underrepresented"*** (lines $\approx$ 310 of the revised manuscript). Further details are found in the revised manuscript. Nevertheless, the close association between 13 km grid spacing with parameterised convection and storm-resolving (e.g. 1 km) grid spacing with explicit convection exists by the virtue of these being the simulation settings in nearly all cases.

Furthermore, we think it is an important factor that the systems in parameterised ICON (and convective systems in most parameterised configurations) seem to mix dynamical behaviour of stratiform and precipitation convective systems in a biased way, leaning more heavily towards the stratiform side, whereas storm-resolving models seem to (over-?)lean towards strongly convective, especially in ICON. This is likely to a large extent explained by grid spacing, although we cannot be certain from our study whether the parameterisation of convection specifically also has a substantial effect (not ruled out by our results; To study this question explicitly, we would ideally need additional grid spacings of 3-10 km in the grey zone with both explicit and parameterised deep convection, which could anyway further clarify on some of the findings from this work.) Therefore, we (slightly over-)phrased it as (variability in convective organisation and geometry) "cannot be represented" in ICON-PAR versus it being represented very explicitly in convection-permitting set-up. In reality, and in rephrasing the statements, we should rather say that the organisation process can be weakly represented in ICON-PAR/13 km grid spacing compared to good representation (possibly even biased under- or overrepresentation of organised updrafts) simulations with ICON-PER/1 km grid spacing.

In addition, to be more specific on the implications for resolving gravity waves, the ICON-PAR grid does not allow to represent gravity wave sources and the full spectrum of relevant gravity waves that the divergence consists of. It may mathematically be considered as a coastline problem in 3D: if we represent an island of 100km^2 on a 10 km grid, its coast line would typically be 40 km. However, if we can represent it more accurately at refined grids, the length may deviate strongly. One could think of gravity wave sources in similar ways, which suggests that their interactions can be represented in much more complex ways at refined grids. Furthermore, as the gravity wave spectrum will extend at sub-grid scales in ICON-PAR, the responses can be more accurately simulated with a setup like ICON-PER. This effect, together with the truncation scales (which will strongly suppress variability at sub-100 km scales in the ICON-PAR) is thought to explain the difference in dispersion between ICON-PAR and ICON-PER, and a corresponding brief explicit discussion included in the revised manuscript.

Furthermore, in line with the closing summary of the previous paragraph, we do not agree that the ellipses are (at least typically) fitted to precipitation fields at scales that are resolvable in the parameterised convection setup: the truncation of the grid scale is typically at wavelengths of about 6-8 grid cells of the grid spacing [e.g. [8]], which translates to about 100 km for ICON-PAR. Everything smaller than this scale is in the grey zone and partially resolved to unresolved, with a rapid decrease of the resolved fraction, presumably especially at 4-8 times grid spacings. The size of large ellipses grows to about 100 km in ICON-PER, but there are plenty of smaller ellipses in our dataset as well. Therefore, we are a little concerned that the reviewer may just be too optimistic about the ability of numerical models to resolve mesoscale processes at lengths less than 8 times grid spacing, which is readressed in the following.

4. *'"192: It is stated that the divergence fields in the PER simulation are low-pass-filtered to remove scales below 45km; is the same low-pass filtering also applied to the divergence in the PAR simulations, to ensure fair comparison? Since PAR has a grid-size of 13km, it may have some variability at scales smaller than 45km, so should be filtered in the same way as the PER simulations? Also, why is it nessecary to low-pass-filter the fields first, given that the analysis only considers the divergence spatially averaged over large boxes anyway? Is there a danger that the low-pass filtering may spread some of the divergence across the box boundaries, so that it is spuriously missed in the analysis?"'* and *'"The method of defining regions of interest and calculating the mean upper- level divergence differs between the 1km explicit versus 13km parameterised simulations, so it is not a like-for-like comparison (the text implies that the 1km run divergence field was first filtered onto a horizontal scale significantly coarser than the resolution of the parameterised run, yet no filtering was applied to the parameterised run for consistency; also much larger boxes were used in the parameterised run)."'*

The filtering operation should in our opinion be separated from the box method. The box method on itself is applied to both ICON-PAR and ICON-PER; in the processing, the filtering is essentially applied to in the first place assure that the convection-permitting ICON dataset does not represent variability that the ICON-PAR simulations cannot represent, after which to equivalent datasets the box method is used for integration.
In retrospective, it might possibly seem a better choice to consider multiple filtering operations, including no filtering, and compare a couple of those options between the two simulation methods as well. On the other hand, this would have created a very comprehensive study and data processing procedure, probably a too comprehensive one for just one manuscript (meaning that two to three manuscripts could be needed to disseminate the results).
Nevertheless, the filtering procedure ought to serve the purpose of creating comparable datasets in simulations of different resolution, rather than diminish comparability (as possibly suggested by the reviewer). This is because the operation should, when applied to ICON at 13 km grid spacing, filter out very little variability: the dynamical spectra of a numerical model drops rapidly at length scales of about 6-8 times grid spacing [[8], and also many other studies, which investigate spectra of model simulated dynamics]. At the wavelengths of up to 45 km, the 13 km resolution ICON simulations can be assumed to be heavily truncated and this range falls well below the effective resolution (which is somewhere in the range of 60-120 km)! The scale at which most of the variability gets truncated is just below this effective resolution, in the range of 40-60 km wavelengths. To emphasise why we chose to filter accordingly, we have expanded the paragraph on filtering with a more specific statement in the revised manuscript: ***"The filtering step assures that the box integrations that we carry out are applied to datasets with very similar truncation scales."*** **(lines 204-205)**
The choice of using a low-pass filter was further motivated by the different grid spacings of ICON-PAR/ICON-PER and the ability to track mesoscale divergence at upper levels within the ensemble members of ICON-PER, without being contaminated by high-amplitude small scale signals of small-scale (5-20 km) divergence patches (associated with convection and gravity waves at those scales).
We have selected a data processing procedure and are confident that it may slightly affect individual samples in mostly random directions (e.g., as the reviewer mentions, by incidently moving some random divergence just to the other side of box boundaries). The assessed conditional statistical signals, however, will not be affected by such processing issues that average out over the lifetime of the systems and, even further, over the full dataset.
The box boundaries have been inserted with care, but because of events such as merging of convective systems and variations in storm tracks, it can never be assured that individual samples in such a case study are "well-behaved" (as controllable or predictable as in idealised simulations, independently of data processing and filters). The statistical assessment has nevertheless allowed us to assess uncertainty in the findings.

5. *'"298 - 305: I am struggling to understand this paragraph (describing the identification of the convective systems in the PAR simulations); consider clarifying? It says that "corresponding precipitation moves together with conditionally unstable or lifted air masses" and "A typical (mesoscale) convective system is easily contained within a box of several to tens of grid cells in each horizontal direction for ICON-PAR" This seems to say that the PAR simulation should produce well-resolved regions of precipitation which move with the flow, but then the text says that only 3 static boxes are used to define the convective systems in PAR. Why not define moving boxes to track the systems, in exactly the same way as is done for the PER simulations?"'*

   We could have made the choice to further investigate the systems in ICON-PAR. However, this has not been done, as from each system, we have analysed several ensemble members (and physically perturbed simulations) to make sure that the statistical sample size allows for safe conclusions. Furthermore, there are fewer convective systems in these ICON-PAR simulations - the three mesoscale systems over Central Europe have all been analysed, where essentially the Central-Germany system is a child of the Alps System. A workflow for data processing equivalent to ICON-PER is not beneficial if only a small set of systems exists and the convection scheme makes the cloud system intrinsically a constant regeneration of itself.
   It is possible to analyse the systems over a longer duration, but for the family Alps/Central-Germany, the intensity (rain rate) lowered in an intermediate stage. The strongest diabatic signals from convection will occur when the system is most active, and we include at least one record at low precipitation rate (when one system did not exist for a certain simulation).
   Last, but not least, the ICON-PER simulations have this interesting scatter in their ratio D, which ICON-PAR does not have. Therefore, investigating residual variability and increasing the sample size

was clearly not prioritised over the corresponding more detailed analysis of ICON-PER.

6. *'"The divergence profiles shown in figure 6 do not vertically-integrate to zero for the 1km simulations, implying there is strong net convergence of mass into the defined boxes (see my point 6 below for detail). Whilst the model may solve the compressible equations, the magnitude of this mass imbalance does not seem realistic, which makes me worry that there has been mistake in the analysis?"'* and *'"Figure 6: Do the profiles of ensemble-mean box-mean divergence satisfy mass continuity?? In ICON-PER, it looks as though the vertically-integrated lower-troposheric convergence is significantly greater than the upper-tropospheric divergence. Assuming the rate of change in the mass of air contained in the box is small, we should have:*

$$\int_0^{z_{top}} \rho(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y})dz = 0 \tag{1}$$

and assuming small deviation from hydrostatic balance ($dp = gdz$):

$$\int_{p_{surf}}^{p_{top}} (\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y})dp = 0 \tag{2}$$

But if I try to vertically-integrate the divergence in figure 6 over the y-axis pressure coordinate "by eye", it does not balance and there is strong net convergence into the box (especially at earlier times). e.g. considering the time-averaged profile for PER (fig 6b, black line), we have convergence between 900 and 400 hPa averaging around $0.4 \times 10^{-4} s^{-1}$, *giving:* 5×10^4 *Pa* $\times 4 \times 10^{-}5$ $s^{-1}/g \approx 0.2$ *kg* m^{-2} s^{-1} *And we have divergence between 400 and 180 hPa averaging around* $0.5 \times 10^{-4} s^{-1}$, *giving:* 2.2×10^4 *Pa* $\times 5 \times 10^{-5}$ $s^1/g \approx 0.1$ *kg* m^2 s^1 *So the convergence and divergence seem to be out of balance by about a factor of 2. The mass imbalance of the order 0.1 kg* m^2 s^1 *would increase the mass of air in the box by 25% during the 7 hour period shown, which is clearly not realistic. I can't see this imbalance in the divergence profiles for the PAR simulation (fig 6b, green line, has similar upper tropospheric divergence to PER, but much weaker lower-tropospheric convergence, so looks in-balance). Similar for the PAR divergence profiles in figure S5. Does this suggest there is something wrong, either with the model or the analysis method in PER? Could the discrepency in PER be somehow due to the low-pass filtering?"'*
There is indeed a mistake in Figure 6, where the ICON-PER data display mass divergence data rather than divergence data. Therefore, Figure 6 has been corrected accordingly in the updated manuscript, leading to balanced profiles.
This issue has not been found in Figure S5, and Figures 6b and S5 should display similar divergence profiles for ICON-PAR, accounting for the respective pressure axis, which means they are at/very close to mass balance.
We thank the reviewer for the close and attentive inspection of the figures, which allowed us to correct this inconsistency in Figure 6!

7. *'"Comparing figure 4 with figure 7, it appears that the boxes used to define the convective systems (and compute spatially-averaged divergence and precip rate) have quite different size in PER and PAR, so it is not a fair comparison? The example box in PER shown in figure 4 appears to be a square with side length about 1 degree in latitude (110 km), whereas the boxes for PAR shown in figure 7 are rectangles with side lengths 2-4 degrees. The use of systematically larger boxes in PAR might risk including more of the background subsidence within the box, so that the spatially-integrated divergence is reduced?"'*
The size of the boxes could unfortunately not be chosen smaller than 2-4 degrees in the ICON-PAR; the convective systems inevitably grow larger in this simulation. It is a process we cannot control, and since we have a large dataset for this case, we could gladly profit from simulations of the same day and event. The box size in ICON-PAR could be considered a small caveat of the case study.
In the discussion of [7], we have described the impact of the box size on the analysis; of course we cannot rule out that subsidence takes place over the chosen (larger) boxes in ICON at 13 km grid spacing, which may not always be completely realistic; this is by construction implied in the structure of most frequently used convective parameterisations, even if it may not always be a realistic representation of mesoscale convective systems.
Nevertheless, carefully studying the divergence at upper levels, the convective systems and precipitation rates in ICON-PAR, we have assured that locations with variability in divergence are included in our analysis. This strategy maximises the reliability of the box analysis at the minimum of "cost" of included remote subsidence (based on the discussion of [7]).
The trade-off of a highly dynamical case is the displacement of some of the upper tropospheric divergence variability, by strong upper-level winds, to the northwestern flank of the convective systems,

which inevitably resulted in the box size we used in ICON-PAR. As opposed to an idealised case study, where the storm-relative winds are controllable, this inevitably is the best we can do in the current study (at least for a mid-latitude case where organisation is predominantly controlled by wind shear).

8. *'"393 / figure 8: The text says that "from here on vertically integrated values (of upper-level mass divergence) are used". But the quantity plotted on the y-axis says it has units kg m3 s1, which is presumably just $\rho(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y})$ Shouldn't the units be kg m2 s1 if this has been vertically-integrated? Also the lines shown in the figure for comparing PAR with PER (and the discussion around L435) indicate a correction to the divergence to account for the different thicknesses of the outflow layers, but this should not be needed if vertical integrals were plotted as suggested in the text. I would recommend calculating the vertical integral of mass divergence for both PER and PAR and comparing those:*

$$\int_0^{z_{top}} L(z)\rho(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y})dz \tag{3}$$

where L(z) is 1 in the outflow layer, zero at other heights. The result then has units kg m2 s1, which can be compared more directly with precipitation rate (which can be expressed in the same units!) Also Re the discussion around L444 (where the text suggests that the apparent reduction in divergence-precipitation ratio for the 13km simulation with only shallow convection parameterised is due to the outflow layer extending lower-down); this could easily be remedied by plotting the above vertically-integrated divergence, and setting the outflow-layer mask L(z) interactively so that it always captures all model-levels within the time-varying outflow layer depth (instead of setting L(z) to zero below a fixed height)."'
Essentially, the above approach is equivalent to what has been done for this work. Vertical integration of divergence over the whole layer depth causes downdraft divergence at lower levels to be included (which is partially visible as mean feature towards the end of the time series in Figure 6). By setting a mask this issue is resolved, as the reviewer states. Exactly this has been done. Nevertheless, substantial divergence features from downdrafts do vertically overlap with outflow from shallow convection in ICON-PER, which means that the perfect vertical mask does not exist.
Whether the processing leads to a divergence in kilogram per cubic or squared meter depends on whether one concludes by dividing the whole integral (above, equation 3, as inserted by the reviewer) by length scale L or not, to obtain the mean value of the integrand. Both are leading to exactly the same result [while the way it is currently plotted, one can also roughly intercompare the local intensity variation of outflows (integrand), which means it is purely a matter of taste, which of the two is favored]. The manuscript has been corrected accordingly.

9. *'"L537, L580, section 7.2.2 (relation between CMT and outflow strength); The study finds a positive correlation between momentum flux and upper- level divergence (with both normalised by precipitation rate). From this, they conclude that convective momentum transport has some impact on the upper-level divergence. However, correlation does not imply causality. It might be that upper-level divergence (or some aspect of storm structure which is correlated with it) has modified the convective momentum transport? e.g. more organised convective systems may have quite different pressure gradient forces acting on the horizontal momentum of the updrafts, and this will influence the CMT"'*
This is a very interesting comment; there is indeed no complete certainty about causality, based on the correlation signal.
We agree that both pressure perturbations and momentum/accelerations by turbulent motion can affect the upper tropospheric flow, next to the dominant cause of divergence - the heating in the middle troposphere, which partially arises from pressure perturbations.
Now, pressure perturbations are resolved differently between say, the simulation at 13 km grid versus the other at 1 km grid spacing - which is a result of differences in the gravity wave spectra that can be represented. Hence, our view is that the role of pressure perturbations for CMT have been interpreted (in existing literature, e.g. [9]) in a way that is not very clear to us. It might (as the reviewer suggests) imply a relation of CMT to storm structure, but we could not discern a relation from the current study as a consequence of the chosen methodology, and neither from [7] as a consequence of no relation between CMT and upper tropospheric divergence rates - CMT is correlated to convective organisation, neither to storm structure nor to upper-level divergence in that work. Furthermore, which part of the gravity wave spectra is resolved is affected by the numerical model and its resolution; hence the pressure perturbations that excite the resolved momentum acceleration depends on resolution by definition. Now to us, the question arise, what do we define as

- CMT momentum tendencies

- Divergence response
- Pressure perturbations

Since pressure perturbations are described to (at least partially) drive both CMT tendencies and divergence - [9] and our work. As the grid refines, we will generally resolve mesoscale convective pressure perturbations better and could re-attribute some momentum tendencies to the pressure fluctuations around mesoscale convective systems.
We would argue however, that pressure perturbations are also largely a result of the localised heating source, and they partially propagate as gravity waves (which affect pressure themselves).
So it is fair to ask the question: which part of the divergence is aliased onto a process classically called convective momentum transport and arising from sub-grid turbulence at 1 km and finer grid spacings and which part is driven by heating; which part is driven by the convective updrafts and how? Or, at any grid spacing, which part of it is arising from sub-grid turbulence and which part is heating-induced? Is much of what we are often calling CMT essentially sub-grid dipoles of divergence/convergence, arising from gravity wave interactions, that can propagate into the surroundings of updrafts and downdrafts? This is a question well beyond the scope of this work and ultimately also more closely connected to turbulence theory than to deep convection.
Nevertheless, this more fundamental question could challenge the views on CMT, but does not imply that one may not be able to fully differentiate between the two (at least in our view) - given that the differentation is a function of grid spacing. Therefore, in our view, it cannot be stated with certainty that the causality question does actually matter, nor that the causality may indeed be wrongly described in our study.
We would say it is an interesting philosophical question to address in a separate work, one day (about which reviewers or readers would be welcome to exchange thoughts).
Based on the reviewer's suggestion, we have chosen for an even safer formulation than before to describe the relation between CMT and the found correlation, avoiding the suggestion of proven causality in the updated manuscript completely, by correcting *"Given a certain precipitation rate, it has been found here that flow perturbations induced as a result of convective momentum transport have some slight impact the mass divergence slightly in ICON-PER. "* to ***"Given a certain precipitation rate, it is suggested by the findings here that flow perturbations induced as a result of convective momentum transport can likely impact the mass divergence rate slightly in ICON-PER"*** (lines 605-606). However, the earlier formulations already did not seem to suggest proven causality in our opinion (e.g. choice of "probably" rather than certainly).

10. *'"Minor corrections:*

- *L59: missing reference (?) - maybe just need to rerun bibtex and latex?)*
- *L247: typo "quantity" should be "quantify"?*
- *L397: Typo "symbolsIn" (missing full-stop and space between sentences).*

"'

Correct; consistently with replies to the other reviewer, these have been corrected for the revisions.

- *'"L230: Seems to say that u'w' and v'w' are vertically-integrated up to model-level 25 to obtain the vertical integral of CMT acceleration up to this height? Is this correct? Surely CMT acceleration is obtained by vertically differentiating the eddy fluxes u'w' and v'w' ? Therefore, the vertical integral of CMT acceleration below level 25 is just the value of u'w', v'w' at level 25, and no vertical integration is needed? (table A1 seems to say that the analysis does indeed just use the mean values at level 25; consider clarifying this in the text?) "'*
  The analysis of the reviewer is correct; We have clarified the text further (L246-247), regarding this point.
- *'"L477: says "Figure 9 suggests that low A generally corresponds with low D at high precipitation intensities (6 mm/h)". However, nearly all the data-points in figure 9 with precip-rate 6 mm/h are in the same bin for A (0.45), so any dependence of D on A in this region of the plot cannot be seen by the reader?*

"'

Indeed, we will correct the text by removing the practically useless statement here.

References

[1] Edward Groot, Patrick Kuntze, Annette Miltenberger, and Holger Tost. Divergent convective outflow in icon deep convection-permitting and parameterised deep convection simulations. [original pre-print and revised version]. *EGUsphere*, 2023:1–32, 2023.

[2] The two referee comments of referee 1 on ”divergent convective outflow in icon deep convection-permitting and parameterised deep convection simulations.”, 2023.

[3] Referee comment 3 on ”divergent convective outflow in icon deep convection-permitting and parameterised deep convection simulations.”, 2024.

[4] Referee comment 2 on ”divergent convective outflow in icon deep convection-permitting and parameterised deep convection simulations.”, 2023.

[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] L. Bierdel, T. Selz, and G. C. Craig. Theoretical aspects of upscale error growth on the mesoscales: Idealized numerical simulations. *Quarterly Journal of the Royal Meteorological Society*, 144(712):682–694, April 2018.

[7] Edward Groot and Holger Tost. Divergent convective outflow in large-eddy simulations. *Atmospheric Chemistry and Physics*, 23:6065–6081, 2023.

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

[9] Robert A. Houze. Mesoscale convective systems. *Reviews of Geophysics*, 42(4), 2004.