WCD manuscript ID egusphere-2026-5124

The role of grid resolution, latent heat release, and initialisation time on the simulation of Storm Éowyn and its impacts

Seraphine Hauser, Lukas Papritz, and Heini Wernli

Final author comments

We would like to thank all three reviewers for their detailed, constructive reviews and the positive, insightful comments on the manuscript. During this ongoing revision stage, we here provide first replies to the main content-related comments from all reviewers. Given some overlaps in the comments and to avoid redundancy, we have grouped together related points where appropriate. We already mention here that we plan to implement the following major changes in the revised manuscript:

- Streamline the paper and reduce information about secondary aspects of Storm Éowyn that are not essential for the key points of the paper. Also, we plan to shorten the storm overview section (Sect. 3) and to remove Fig. 8.
- Include results about additional simulations (see comment 1).
- Clarify the meaning of the DDC (diabatic deepening contribution).
- Improve the investigation of the track sensitivity to latent heating.

Comments by the reviewers are in black and our replies are shown in **this color**.

(1) Role of grid spacing/horizontal resolution and convection treatment

Reviewer 1: The resolution experiment does not necessarily isolate the effect of horizontal grid spacing. Convection is parameterized at 50, 20, and 13 km and disabled at 7 and 3 km. Time step, numerical diffusion, effective resolution, resolved orography, and the partitioning between resolved and parameterized vertical transport also change. The manuscript nonetheless attributes most differences to horizontal resolution and later concludes that explicitly resolving convective-scale processes improves cyclone intensity. The non-monotonic changes at 13 and 7 km could arise from the switch in convection representation rather than grid spacing. The 13-km run is described as being in the convective grey zone, but 7 km is also within this transition for deep convection. A 3-km grid is convection permitting, not convection resolving. The experiment cannot readily distinguish these effects. I encourage the authors to perform matched convection-on and convection-off simulations at 7 and 13 km. A scale-aware convection sensitivity would be preferable if available. If these experiments cannot be added, the manuscript should consistently describe the results as sensitivity to model configuration rather than just resolution.

This is a really good and valid point and we are thankful for the reviewer's comment. For the revisions, we have created additional experiments, such that for all horizontal resolutions considered (3, 7, 13, 20, 50 km), we can compare how convection treatment affects the results in terms of the evolution of SLP_{min}, the cyclone track, as well as the cyclone structure and associated surface weather. For the 5 different grid spacings, we now have three different experiments each: (i) convection represented explicitly (OFF), (ii) only shallow convection parameterized (SHAL), and (iii) deep and shallow convection parameterized (ON). See Table 1 below for an overview of the experiments we will evaluate. With this, we plan to restructure Section 4 in the paper and rename it to "Sensitivity of cyclone characteristics to horizontal resolution and representation of convection".

Table 1. Overview of all ICON simulation setups used in this study.

experiment abbrev.	approx. grid spacing (km)	ICON grid	time step (s)	regular grid (°lat/lon)	parameterised deep convection	parameterised shallow convection
3k_OFF	3	R3B9	20	0.025	off	off
3k_SHAL	3	R3B9	20	0.025	off	on
3k_ON	3	R3B9	20	0.025	on	on
7k_OFF	7	R3B8	35	0.05	off	off
7k_SHAL	7	R3B8	35	0.05	off	on
7k_ON	7	R3B8	35	0.05	on	on
13k_OFF	13	R3B7	70	0.125	off	off
13k_SHAL	13	R3B7	70	0.125	off	on
13k_ON	13	R3B7	70	0.125	on	on
20k_OFF	20	R2B7	90	0.25	off	off
20k_SHAL	20	R2B7	90	0.25	off	on
20k_ON	20	R2B7	90	0.25	on	on
50k_OFF	20	R3B5	200	0.5	off	off
50k_SHAL	20	R3B5	200	0.5	off	on
50k_ON	50	R3B5	200	0.5	on	on

(2) Methods: Design of "no latent heating" experiments and cyclone tracking

(a) Reviewer 1: The text alternates between stating that the latent-heat constant is scaled and that temperature tendencies associated with phase changes are scaled. These are not necessarily the same in terms of the numerical model modifications. Please state exactly what was changed in the source code. In particular, clarify whether alpha multiplies the latent-heat constants, the diagnosed temperature tendencies, the underlying phase-change rates, or both thermal and moisture tendencies. Also please state whether positive heating and negative cooling tendencies are both scaled and document the consequences for total-water and moist-energy conservation.

We thank the reviewer for this comment. To provide a short answer already here: We modify the temperature tendencies that result from phase changes. This way both positive temperature tendencies from heating (condensation, freezing, deposition) and negative temperature tendencies from cooling (evaporation, melting, sublimation) are scaled. In the revised version of the manuscript, we will include a better explanation of our "suppressed/no LH" experiment design with more details on what is suppressed and what stays untouched in the ICON code. With this, we hope to clarify any remaining ambiguities.

(b) Reviewer 1: Regarding the cyclone tracking algorithm, the search boxes in degrees are latitude dependent. Refer to distances in kilometres.

The cyclone tracking algorithm has been adjusted and updated, and is now working with boxes/radii expressed in kilometres. The tracks of Éowyn in all simulations have been regenerated with this new version.

(3) Budget of diabatic deepening contribution

Reviewer 1: The diabatic deepening contribution is not a process budget. It is defined as a normalized difference between the minimum pressures from two nonlinear simulations, which can be at different times and locations. It is sensitive to forecast length, track divergence, land interaction, lateral-boundary history, and the pressure at initialisation. It does not partition cyclone deepening into additive dry and diabatic components. This is particularly important for the range of 12–53% discussed in the manuscript. Later initialisations have a smaller denominator because part of the cyclone deepening has already occurred. This can produce a large percentage even when the absolute control–no-LH difference is not that large. Early initialisations allow more time for nonlinear trajectory divergence. The U-shaped dependence may therefore be partly a property of the metric rather than the physical importance of latent heating changes. It would be helpful to discuss this point more and compare the results with a common verification interval to assess how important of an issue this is. Pressure differences could be evaluated at matched valid times over a common verification period. These comparisons would show how much of the U-shaped behavior arises from normalization and differing integration lengths. Unless more support is provided (which could include a budget analysis), "diabatic sensitivity" would be more accurate than "diabatic contribution."

We thank the reviewer for this thoughtful comment. The DDC (diabatic deepening contribution) is a quantity we calculated to determine the role of diabatic heating to cyclone intensification. As you mention, this is a very simple and pragmatic measure to quantify the role of diabatic processes, introduced in the review paper by Wernli and Gray (2024). It mainly serves to provide an overall estimate of how much deeper the cyclone became due to the additional deepening associated with diabatic processes and it enables a simple comparison across case studies. We realize that applying this measure to simulations that start during the deepening phase (as we did in the original manuscript) is confusing. In the revised manuscript, we will only use the DDC for the simulations started before cyclogenesis, i.e., before 12 UTC on 22 January 2025, which cover the full deepening phase of the storm.

(4) Significance of correlation results in Figure 8

(a) Reviewer 1: The hourly samples in Figure 8 are taken from the same evolving cyclone and are correlated. Their significance cannot be assessed with a standard correlation test that treats each hour as independent. The dominant cyclone life-cycle trend may explain much of the relationship. Both low-level PV and strong-wind area respond to cyclone intensity, pressure-gradient strength, frontal contraction, storm size, stratification, translation speed, and boundary-layer fluxes. Their correlation does not show that larger low-level PV causes a larger wind area. A small increase in correlation from the 50-km to the 3-km run doesn't establish that higher resolution improves the simulated causal relationship. Because the hourly samples are serially correlated, the correlation analysis should account for autocorrelation, for example through an effective sample size or a bootstrap approach. If the authors want to assess whether the relationship extends beyond the cyclone life-cycle trend, partial correlations controlling for intensity, storm size, and time would be useful here. Otherwise, Sections 4.3 and 5.4 should describe the result as an association rather than causations.

We agree that the hourly values of low-level PV and area of high low-level wind speeds in Figure 8 come from one cyclone and are strongly correlated. When changing the structure in Section 4 and discussing not only the effects of horizontal resolution but also the sensitivity to convection treatment, we will remove Figure 8 and the associated discussion from the manuscript.

(b) Reviewer 3: In Figure 8 and the discussion of it in section 4.3, why perform the correlation calculation using all the data shown? It seems clear from the figure that the linear correlation between low-level PVU and high-wind coverage does not hold early in the cyclone's life cycle. I'd also wonder whether the trend that "[o]verall, higher resolution yields stronger correlations" (line 304) if the early-life-cycle points are removed.

Thanks for this comment. While we agree that the correlation would be higher if the early time steps from the life cycle were removed, performing the correlation analysis over a subjectively determined time period would feel "designed" in order to get a higher correlation. Referring to the reply to the previous comment we decided to not show this analysis anymore in the revised version of the manuscript.

(5) Track sensitivity to latent heating

(a) Reviewer 1: The no-LH simulation develops a slower and more poleward track, but the mechanism is not conclusively demonstrated. Cyclone intensity, low-level circulation, divergent outflow, jet structure, downstream ridge amplification, steering flow, and propagation all evolve together. Figures 11 and 12 document differences between the simulations, but they do not establish the mechanism described in the abstract and conclusions. The accuracy of this QG diagnostic for a strongly diabatic, rapidly deepening cyclone with substantial curvature and ageostrophic flow needs to be assessed. The manuscript does not compare the diagnosed propagation direction or speed with the actual cyclone-motion vector, as I understand. It also does not quantify the residual error. A 24-h cyclone-relative composite may make the temporal

ordering of the response difficult to interpret. Please compare the observed cyclone-motion vector with the direction from the cyclone center toward the diagnosed maximum cyclonic development or propagation forcing. Please note whether the magnitude and location of the forcing are consistent with the observed propagation speed and direction through time.

Thank you for this comment. We agree that a more detailed analysis of the cyclone propagation is necessary, in particular to what extent the thermal wind and the “dry” QG diagnostic can capture the actual propagation of the diabatic cyclone.

(b) Reviewer 3: One of the central claims of this section of the paper is that reducing the diabatic forcings from latent heating yields a northwestward shift of the cyclone track (line 598), which contradicts idealized GCM studies on latent heating and cyclone propagation. However, I’m not sure exactly how robust these results are from the data that are shown. In Figure 10(b), for instance, it’s true that the no-LH cyclones lie poleward of the control-run cyclones at 23 Jan and 24 Jan, but it seems that by 25 Jan there’s very little systematic difference between the latitudes of the two sets of cyclones? Since the no-LH tracks are biased poleward at initialisation in this figure, this would imply that the control-run cyclones actually propagate northward (and eastward) faster to ‘‘catch up’’ to the no-LH ones, consistent with the earlier idealized studies (and that the bigger question is why the no-LH cyclones are more poleward at initialisation in the first place). These observations are just from inspection of the figure; if the enhanced northward propagation of the no-LH cyclones is actually (statistically) significant, it would be helpful to show this more systematically. There are similar ambiguities in Figure 11—for instance, in subfigure (e), it seems like the control and no-LH cyclones are at the same latitude on 23 Jan 12 UTC.

Thanks for your helpful comment. In the revised version, we will present an analysis on why in the very early phase of the cyclone (around genesis) these jumps in the cyclone track occur and will show that these are not necessarily linked to “artifacts” of the cyclone tracking algorithm. Additionally, we will perform a detailed analysis of the cyclone track differences between the control and no-LH experiment and will include this in the revised version of the manuscript. With first results at hand, we can confirm your hypothesis: Across all initial times, the no-LH cyclones do not propagate northward faster. The statistically solid differences are that no-LH cyclones move more slowly eastward and they are weaker in intensity.

(6) Using ERA5 as “truth” for verification of ICON simulations

Reviewer 1: ERA5 is used as truth for the extreme cyclone. Please provide some verification references.

Thank you for this valuable suggestion. ERA5 reanalysis is used as the evaluation baseline in this study, given its high temporal resolution compared to, for example, the IFS analysis from ECMWF (available only every 6 hours). To verify the accuracy of ERA5 for this specific storm, we will compare ERA5 data against observational weather station/buoy data across and around

the British Isles during the cyclone passage and will comment on our outcome in the point-by-point reply submitted with the revised manuscript.

(7) Separating low-level dynamics from surface impacts

Reviewer 1: This section discusses near-surface winds but analyzes 850- and 925-hPa winds. Separate low-level dynamics from surface impacts.

We agree that throughout the manuscript we present too many different proxies for low-level wind speeds, such as 850 hPa wind speed, 925 hPa wind speed, 10 m wind speed, and 10 m wind gusts. In the revised version, we will stick to a maximum of two low-level wind variables, most likely to 10 m wind speed, which is very similar to the lowest model level wind speed in ICON, and wind gusts to compare to the observed gusts around the British Isles.

(8) Relative role of numerics-caused PV mixing for resolution experiments

Reviewer 2: I basically agree with the conclusions of the authors obtained from the resolution simulations, particularly about the PV tower interpretation for the high-resolution experiments, though the relative role of numerical diffusion and physical processes is not fully addressed. My suggestion is that the authors may analyze (or mention only if detailed analysis is not possible during the revision) the numerics-caused PV mixing (PV homogenization) which may lead to the weakening of PV, PV-gradient, and jets in the coarser-resolution simulations.

We thank the reviewer for the careful reading and for the supportive assessment of our conclusions, particularly the PV-tower interpretation of the high-resolution experiments. We agree that the relative roles of numerical diffusion and physical processes are not explicitly considered in our study. Numerically induced PV mixing could indeed contribute to some extent to the weaker PV, PV gradients, and jets in the coarser-resolution runs.

However, explicitly quantifying the numerical diffusion tendencies is not possible and therefore we can not present such an analysis. Still, we aim to include the point of numerical diffusion as limitation in the revised manuscript when discussing the vertical PV cross sections.

(9) Role of latent heat release on PV gradient

Reviewer 2: About the simulations about the role of latent heat release, is it possible for the authors to take a look at the role of latent heat release on the strength and positioning of PV gradient (and also the orientation of upper- and mid-tropospheric jet)?

Similar to the comment above, we regard this as an interesting suggestion, but one that goes beyond the scope of our paper. Bukenberger et al. (2023) investigated in detail the processes affecting upper-level PV gradients in km-scale simulations. What concerns our study, we prefer not to add this aspect to our study (which is already rather extensive), but we will comment briefly on potential differences in the upper-level jet in the revised version of the paper.

(10) Initialisation time analysis

Reviewer 3: The sensitivity of the forecast skill of this storm to initialisation time is a rich and complex subject for inquiry, but as currently written, Section 6 mainly seems to describe the results and emphasize conclusions which are already common knowledge. For instance: ‘‘A correlation exists between the SLP*min deviation (ICON vs. ERA5) and the initialisation time, demonstrating that later initialisations yield a more accurate representation of cyclone intensity across various horizontal resolutions’’ (lines 613–5). This is just the general nature of forecast error growth, as formally demonstrated by, e.g., Lorenz (Tellus, 1982), and for the Lagrangian statistics of extratropical cyclones by Froude et al. (Mon. Wea. Rev., 2007). Similarly, that increased horizontal resolution improves the skill of short-range forecasts is well-attested in the literature (Buizza, Quart. J. Roy. Meteor. Soc., 2010; Davis et al., Wea. Forecasting 2010; Lei & Whitaker, JAMES 2017).

The discussion in section 6 highlights the non-monotonicity of this correlation, with anomalously poor forecasts occurring on 18 January and 20 January (18 UTC) for instance. On the origins of this monotonicity, the paper says that ‘‘[t]he generally poor performance at these initialisation times is suggested to be attributed to data assimilation errors that yield incorrect initial conditions’’ (lines 564–5). Can you demonstrate this more concretely, for example by considering the nature of the error growth in these particular forecasts? Being more specific on this point would help strengthen this section considerably, I think. (Though, given the complexity of the subject, this may turn out to be an analysis for another paper!)

There are a number of similar passages in this section that have a similarly speculative tone that I would rather see supported by more rigorous data analysis or reasoning: ‘‘The stronger wind gusts...model resolution’’ (lines 501–2); ‘‘This suggests that at coarser model resolutions...prior to cyclogenesis’’ (542–4), etc.

On a similar note, there are some interesting properties of the data which I feel are under-discussed in the text. For instance, why did so many IFS ensemble members fail to form a cyclone on 18 January? Why do the ICON simulations systematically underestimate the cyclone pressure anomaly relative to ERA5/IFS from 19 Jan on (Figure 14, left panel)—is this a resolution effect? Also, as shown in Figure D1, the position biases seem to be of similar direction across model resolutions and shift somewhat coherently with lead time (e.g. from a mostly southern bias on 17 Jan to a north-northwestern bias on 20 Jan). Any thoughts as to why this is the case? Can we learn anything from this about biases in the model or the data assimilation? I bring these points up merely as additional examples of observations where I think this paper could develop some interesting physical insights.

We thank the reviewer for this comment and we will reduce the speculative remarks in this section. Given the many aspects covered in the manuscript, we currently consider skipping Section 6.3 entirely and replacing it with one more analysis on why the ICON simulation and IFS ensembles on 18 January fail to produce a cyclone – therefore putting more focus on current Figure 14 while omitting Figure 15.

(11) Length of the manuscript and superfluous details

(a) Reviewer 3: Section 3 of the paper provides a thorough review of Storm Éowyn's meteorological history. This review is well-written and interesting in its own right, but some of the discussion is extraneous to the rest of the paper and may be excised. For instance, the paragraph discussing the storm's status as a diabatic Rossby wave (lines 180–92) feels parenthetical, and I'm not convinced it is entirely necessary. Similarly, Figures 3 and 5 are very information-dense and partially redundant with each other and with later figures in the text: e.g., both Figures 3 and 5 contain PV maps for 23 January at 00 UTC and illustrate similar meteorological features to those in figures 7 (vertical structure of the PV tower) and 11 (ridge building as the cyclone intensifies).

I would consider shortening this section to emphasize the meteorological features that are important for the later discussion of forecast sensitivity, and to highlight that importance explicitly. This might also allow you to make later sections more concise. I'd also consider eliminating Figure 3, since this feels the least significant to the later sections, and incorporating any especially important information therein into a revised Figure 5.

In the revised version of the manuscript, we will shorten Section 3 and reduce the number of figures from 4 to 2. We agree that there is redundancy in the four figures and information provided, which is not needed for the further analyses in the paper, such as for example the sting jet discussion.

(b) Reviewer 3: Section 5.4 can also be cut and its contents redistributed to other sections. I do like the central figure (Figure 13), but I'm not convinced it deserves an entire section to itself; its interpretation seems like a fairly straightforward consequence of the shifts in the cyclone track and intensity discussed previously. Moreover, the framing of this section strikes me as an empty hypothetical—if atmospheric latent heating suddenly turned off, a lot more would change than this specific storm! (This may be more a matter of taste though.)

We see the reviewer's point and consider moving this figure to one of the other subsections of Section 5. With some changes in the analyses regarding the track sensitivity and initialization time, it is likely that the current Section 5.4 will disappear.

(c) Reviewer 3: Much of the text of Section 6 reads as a ''blow-by-blow'' description of Figures 14 and 15, and this section would be strengthened if the writing focused more on analyzing the data in these figures (see later for more comments about this section).

As described in Point (10) in this document, we will also shorten/restructure Section 6 of the manuscript and will make sure it does not come across as a "blow-by-blow" description.

(d) Reviewer 3: There are a few other passages scattered throughout the study which feel superfluous, e.g.: ''Although this is not the focus...indicating high confidence'' (lines 231–3); ''Interestingly, the stochastic perturbations...wind gusts'' (lines 503–4). Additionally, the abstract

should probably be shortened by a sentence or two—I have no specific requests for what to cut here, and defer to the authors' judgment about which points they'd like to emphasize.

Thank you for spotting this. In order to shorten the manuscript, we will watch out for these superfluous passages and remove them. Also, we try to slightly shorten the abstract in the revised version.

(12) Grammar, Wording and Figures

We received some comments regarding:

- phrasing/wording/speculating ("significant" although no significance test was performed)
- language/grammar mistakes
- missing latitude-longitude labels in figures and unclear coastlines
- Wrong value extent mentioned in figure captions

While working on the revisions of the manuscript, we will address these points and make sure to improve presentation of our results.

References

Bukenberger, M., Rüdisühli, S. & Schemm, S.(2023) *Jet stream dynamics from a potential vorticity gradient perspective: The method and its application to a kilometre-scale simulation*. *Quarterly Journal of the Royal Meteorological Society*, 149(755), 2409–2432, https://doi.org/10.1002/qj.4513**.**

Wernli, H. & Gray, S. L.(2024) The importance of diabatic processes for the dynamics of synoptic-scale extratropical weather systems – a review, Weather Clim. Dynam., 5, 1299–1408, https://doi.org/10.5194/wcd-5-1299-2024.