

Advancing the Capabilities for Efficient Hurricane-Centric Simulations with the Atmospheric Model ICON

Authors: *Senf, F. and Cremer, R.*

Summary

The authors present an automated workflow toolkit that enables the efficient simulation of tropical cyclones on a sub-kilometre scale. This is achieved by running the ICON model sequentially on multiple overlapping grid segments along the cyclone's expected path. Tailored to the ICON model, the toolkit automates several key tasks, including cyclone tracking, grid generation and the preparation of initial and boundary data, as well as conducting the high-resolution segment-based runs. The toolkit's key components rely on existing software packages, such as *tobac* for cyclone tracking, *DWD ICON Tools* for grid generation, and *CDO* for horizontal remapping.

The workflow toolkit is demonstrated using Hurricane Pauline (2020) with nested high-resolution simulations performed using grid sizes down to 300 m. The authors compare different configurations of the grid segments and conduct performance measurements on two different state-of-the-art HPC systems. They also discuss shortcomings of their approach, focusing on spin-up effects resulting from the remapping and merging of the initial conditions. Based on their findings the authors conclude that, compared to traditional approaches with one fixed limited-area domain, the efficiency gains are significant, reducing resource requirements by a factor of 13-175 depending on the configuration details. They also conclude that the observed spin-up effects are acceptable, given the substantial efficiency gains achieved.

General comments

The manuscript is well written and concise. It makes a valuable contribution to the scientific topic of how to conduct km-scale tropical cyclone simulations at a reasonable cost. The proposed workflow toolkit appears to be scientifically sound in most respects and is clearly described. The authors also effectively demonstrate the efficiency of their approach. The visual presentation of a Lagrangian cyclone analysis prototype is well designed, too.

However, I believe the manuscript requires a more detailed analysis of how the preparation of the initial conditions impacts simulation quality. I am not convinced that the analysis of the spin-up effects presented provides sufficient evidence to conclude that the negative impacts on simulation quality are sufficiently small. I have summarised potential pathways towards an extended model evaluation below. In summary, I recommend that the manuscript be reconsidered for publication once the following major points have been addressed.

Major points

Initialization of the high-resolution segment-based runs

The initial conditions (IC) for segment k are generated by merging the base run data with the output data from segment $k - 1$. Section 3.5 examines the spin-up of several thermodynamic

parameters resulting from that procedure during the first few hours of model integration. Spin-up effects are known to occur, if the driving model climate differs from that of the target model, or if the model climate changes significantly with grid resolution. Apart from spin-up effects, the proposed initialisation procedure may generate a significant amount of spurious sound and gravity waves ('noise') which might degrade the quality of the solution. However, the manuscript is lacking a quantitative analysis of the noise level.

As a quantitative measure of this noise, I recommend looking at time series of the domain-averaged surface pressure tendency (see the ICON model tutorial Section 2.2 for a description). Such time series can be extracted from the ICON log files. Specific questions:

- How does the initial noise level of the high-resolution runs compare to that of your base run? To make a fair comparison you might need to repeat the base run at the resolution of your coarsest segment, and to restrict the analysis to the area covered by the segment. However, a direct comparison of the existing runs may already be sufficient.
- Does a reinitialisation interval of 12h or 24h provide sufficient time for the initial noise to decay? In other words, is the timeseries of the surface pressure tendency a flat curve by the end of the reinitialisation interval, or does it still slope? If it still slopes, your reinitialisation procedure may accumulate noise over successive segment runs, as your IC merging procedure will transfer spurious waves from one segment run to the next.

Alternative initialization procedures

Have the authors explored alternative model initialisation procedures apart from IC merging? Might the Incremental Analysis Update (IAU) procedure be applicable? IAU is an effective filtering technique for spurious high-frequency sound and gravity waves (see, for example, the ICON model tutorial Section 11.3). In ICON, it is applied operationally to add analysis increments provided by the data analysis. If you define the difference between your base run fields and the fields from segment $k - 1$ as the increments (or vice versa), the IAU procedure might be applicable to your problem.

Quantitative error analysis

The manuscript lacks a quantitative analysis of the errors introduced by the segment-based approach. Ideally, the results of the segment-based approach and the classic approach (single large LAM domain) should be evaluated against measurements. Only then, can the question be safely answered, as to whether the efficiency gains achieved justify the acceptance of potentially greater model errors. Admittedly, this is easier said than done.

A pragmatic alternative would be to compare the results of the segment-based approach with those of the classic approach at the same horizontal and vertical resolutions. Using the results of the classic approach as a reference would allow you to evaluate differences in noise levels, deviation of the model results, and the impact of the chosen cross-track width.

Lateral boundary conditions and external parameters

Lateral boundary conditions (LBCs) are an integral part of the presented workflow. However, Section 2.3 'Workflow Schemes and Implementation Details' lacks a description of how LBCs

and external parameter fields (e.g. topography) are treated. While the authors touch this topic in several sections of the manuscript, they provide no concise description. Specific questions:

- Are high-resolution segment-based runs at their boundaries always driven by hourly data taken from the base run, or is an additional post-processing step performed on LBCs similar to the merging step for initial conditions?
- Do you run the Extpar tool on the individual high-resolution segment grids, or do you interpolate the coarse external parameter fields from the base run onto the segment grids? One consequence of the interpolation approach is that the high-resolution runs make use of an overly smooth topography. While I am aware that the focus of your simulations is on the ocean areas, the use of coarse external parameter fields is expected to notably impact hurricane simulations during landfall. Therefore, if this constraint exists in your toolkit, I recommend stating this clearly.

There might exist a few more shortcomings, which should be mentioned explicitly in the manuscript, if applicable:

- One possible shortcoming relates to the land and soil model. Initial condition fields containing tiled land surface information cannot be regridded to different resolutions, because the tile partitioning differs between the source and target grids. It is still possible to activate the tile approach for your segment runs, however a tile cold start will be required (i.e. initialisation from aggregated fields). The same restriction applies to the classic approach of a single large limited-area run such as your base run. However, the problem is amplified for your segment runs, as repeated initialisations are required. Hence, your segment-based approach might restrict the usage of the land and soil tile approach.
- Recently, the consequences of using inappropriate LBCs at the time of model initialization have gained interest within the ICON community. By 'inappropriate' I refer to LBCs that differ from the initial conditions read from the initial conditions file. In such cases, the model simulation is likely to suffer from model biases, proportional to the mismatch between the initial LBCs and the initial conditions. A detailed description of this problem is available on the ICON webpage https://docs.icon-model.org/documentation/buildrun/buildrun_input_data.html. It is likely that your model runs are affected by this problem. The proposed solution is to overwrite the initial LBCs with the initial conditions, by activating the namelist switch `init_latbc_from_fg=.TRUE.` (default `FALSE`). If this switch was deactivated in your model runs, I recommend repeating a subset of your runs to check whether this has any impact on the model results presented in your manuscript.

Minor points

1. **Line 44:** '...and the influence of larger scale processes is not captured adequately.' I would merely argue that the influence of small- to regional-scale processes on large-scale processes is not adequately captured by design. Are the large-scale processes not accounted for by the lateral boundary conditions?

2. Optional suggestion: **Line 88:** ‘...two one-way nested domains with grid spacings of 5 and 2.5 km are employed ...’ The wording ‘two one-way nested domains’ might be a bit confusing for readers from the ICON community. If I understand correctly, your base setup consists of a regional domain and one nested domain (one-way nested). In particular, I propose to avoid using the word ‘nest’ for the base domain (DOM01), and to avoid counting it as a nest. The base domain can be a global or regional domain, which may be equipped with one or more nested domains. You might consider adapting the wording in the text (e.g. L153, L157, L205, etc.).
3. **Line 119:** In your example the distance of the track segments A and D in terms of time is $\Delta t_{reinit} + 2\Delta t_{add} = 18$ h. However, in the text you state that it corresponds to the distance traveled by the hurricane, ‘e.g. from 18 UTC the previous day to 15 UTC’, which corresponds to a time interval of 21 h \neq 18 h. Is there a typo in the text, or am I missing something?
4. IC generation **Line 159-166:** This paragraph might benefit from some rephrasing, in order to point out more clearly that
 - CDO is used for horizontal regridding only (no vertical regridding)
 - vertical regridding (interpolation) is performed by the ICON test run, which reads the horizontally interpolated ICs, and outputs the vertically interpolated IC to file right before the model performs the first integration step (time step zero).

Vertical regridding is mentioned again in Section 3.3 (L275-278). I suggest to discuss it only once.

5. In Section 3.3 ‘Exploration of Different Refinement Strategies’ the resolution of your grid segments (DOM01) is R2B11, which is exactly twice the resolution of DOM02 of your base run (R2B10). I assume that you are not constrained to a factor 2 in resolution jump and that it is possible to select arbitrary horizontal resolutions for DOM01 of the grid segments (within physically reasonable bounds)? In particular, I assume it is equally possible to choose R3B10 or R3B11 for DOM01 of your grid segments in this example? Is my assumption correct? Your statement ‘In principle, it is also possible to start initial refinement from grids finer than the base grid’ (L256) seems to point towards this direction, but it is not fully clear to me whether there exist technical limitations in your workflow or not.
6. **Figure 7b:** Can you explain why the minimum surface pressure seems to undershoot when switching to the next grid segment? I would rather expect the minimum surface pressure to increase, as a consequence of regridding and weighted merging. Both, regridding and merging, will likely diffuse gradients in the prognostic fields which I expect to result in increased minimum surface pressure values.
7. **Line 327:** ‘...the 12h-reinit setup is significantly more favorable in terms of resource requirements ...’. Could you please explain why the 12h-reinit setup requires less computing time than the 24h-reinit setup (see Table 2)? Assuming an identical across-track width, switching from 24h to 12h will result in a domain size that is half the original size. However, the number of grid segments required will double. So, where does the observed speedup for the 12h reinitialisation interval come from?

Editorial

1. Figure 1: Gray arrows are a bit hard to see in print form.
2. Table 1: Strictly speaking the column 'reinit' does not fit, as its content clashes with the labeling of the rows (similar for Table 2). It seems like a pragmatic approach to keep the table compact and, unfortunately, I do not have a better solution.
3. Figure 7: 'The minimum pressure ... as in (b) and (c)'. 'as in (b) and (c)' does not make sense to me. Do you mean 'as in (a)'?
4. There is some inconsistency in the way units are defined in the figures. Axis annotations sometimes use square brackets, while at other places a slash is used followed by units that are sometimes enclosed in round brackets and sometimes not. I suggest the consistent use of one style, such as square brackets for example.
5. Line 313: 'Most of the ice produced is located northwest of the hurricane center.' Isn't most of the ice located northeast rather than northwest in Figure 9?