SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

Reporting year	2024
Project Title:	Forecasting at sub-kilometre resolution with the HARMONIE-AROME model
Computer Project Account:	spiecla2
Principal Investigator(s):	Colm Clancy
Affiliation:	Met Éireann
Name of ECMWF scientist(s) collaborating to the project (if applicable)	N/A
Start date of the project:	1 January 2023
Expected end date:	31 December 2025

Computer resources allocated/used for the current year and the previous one
(if applicable)
Please answer for all project resources

		Previous year		**Current year**	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	20 M	19.6 M	30 M	15.96 M
Data storage capacity	(Gbytes)	-	-	-	-

Summary of project objectives

The objective of this project is to explore configurations for running the HARMONIE-AROME limited-area NWP model at hectometric resolutions, and to assess the potential benefits for forecasting over Ireland.

Summary of problems encountered

No major problems on the ECMWF HPCF were encountered.

Summary of plans for the continuation of the project

The work so far has highlighted the problem of dry regions near the boundaries of the computational domain, with a larger domain showing benefit in specific cases. This will be explored further with longer simulations, comparing with other options such as nesting within an intermediate-resolution simulation.

Work on higher horizontal resolutions will continue. The preliminary experiments showed that simply increasing to 500m does not bring immediate benefit, but better validation and tuning of the model will be required; in particular the new scale-aware convection scheme will be tested.

List of publications/reports from the project with complete references

The 2023 results were described in the following:

Clancy C, Fannon J, Harney E, Kokina T, Whelan E.
High-resolution and Dynamics Experiments at Met Éireann.
ACCORD Newsletter No. 5, March 2024

Clancy C, Fannon J, Harney E, Whelan E.
Boundary and Dynamics Options for HARMONIE-AROME at Hectometric Scale
Internal NWP Note available from Met Éireann

Summary of results

The testing in 2023 focussed on boundary options for running HARMONIE-AROME at 750m resolution over Ireland. The results can be summarised as follows:

- coupling the available hydrometeors and cloud water species is recommended
- when looking at different nesting options, the age of the initiating global boundary files tends to dominate performance
- while using IFSHRES LBC directly shows slightly better scores, spatial rainfall analysis shows a benefit to nesting within an intermediate HARMONIE-AROME forecast
- an alternative would be to increase the domain size, but this obviously comes at a cost

In addition to this work on boundaries, a number of other dynamics topics were investigated:

- the MetCoOp 90-level grid was tested, as an alternative to the Météo France one. While there is some evidence of improvements in terms of fog/low cloud forecasting, we do not get the benefit of improved wind accuracy that comes with the latter.
- the XIDT uncentring option was tested for the stormy period. The effect on extremes in terms of MSLP minima and winds was found to be minimal.

The ACCORD Newsletter describing the above is appended to this report.

All of last year's testing used a 750m cubic grid. Work in 2024 has so far involved exploring the effect of horizontal resolution and grid truncation. A 750m linear grid has the same spectral resolution as a 500m quadratic grid, but at about a third of the computational cost. Initial testing during a stormy winter period showed no benefit in terms of verification scores with the 500m grid: in fact, the biases were worse (Fig. 1). However, these biases may be more an effect of the precision used in the model (Fig. 2), and this is currently under investigation.

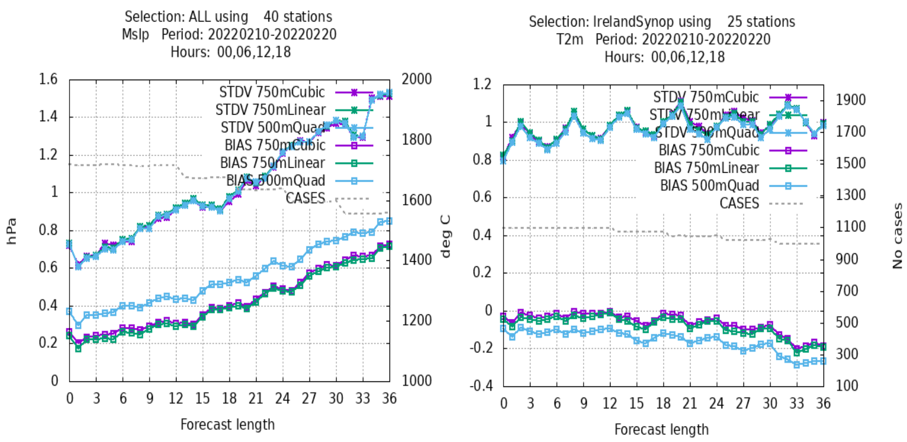

Figure 1: verifcation of MSLP and T2m for 750m and 500m resolutions with different spectral truncations.

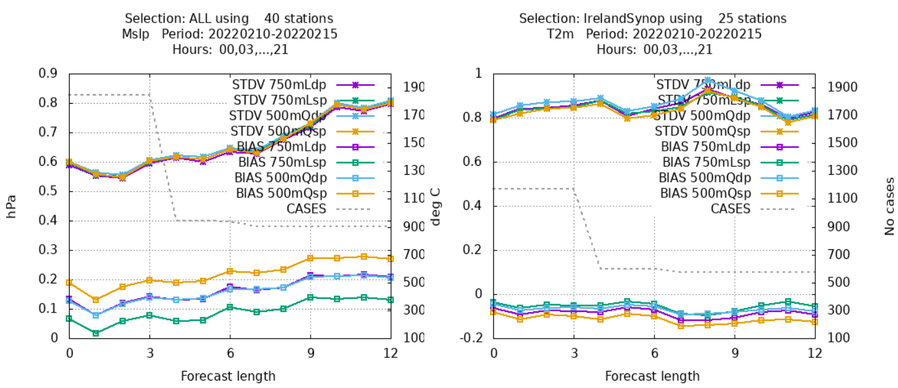

Figure 2: verifcation of MSLP and T2m for 750m and 500m resolutions with different spectral truncations (L linear, Q quadratic) and either single (sp) or double precision (dp).

June 2024

High-resolution and Dynamics Experiments at Met Éireann

Colm Clancy, James Fannon, Eoghan Harney, Tatjana Kokina, Eoin Whelan

1 Introduction

The current operational IREPS suite at Met Éireann uses a horizontal grid-spacing of 2.5 km with 65 levels, similar to other centres running HARMONIE-AROME. Based on previous work (Clancy et al., 2022) investigating the potential benefit of higher resolutions, we are currently running a parallel e-suite called HECTOR, which uses a 750 m grid with the 90 vertical levels used by Météo France (hereafter MF_90). HECTOR uses boundaries directly from IFSHRES; previous experiments found this to be preferable to nesting within an intermediate HARMONIE-AROME host.

Soon Met Éireann operations will be carried out within the UWC-West group, and the common domain will have a 2 km grid with the MF_90 vertical levels. Additionally, a deterministic forecast will run every hour, giving more LBC options for the HECTOR suite. In this article we summarise the key results from a series of experiments to determine the “optimal” LBC configuration (if it exists) for a high-resolution subdomain like HECTOR. A more comprehensive technical note detailing this work is available from Met Éireann.

In addition, a number of dynamics settings were tested, both in these high-resolution and standard configurations during Storm Ophelia in October 2017. The various domains used throughout are shown in Fig. 1.

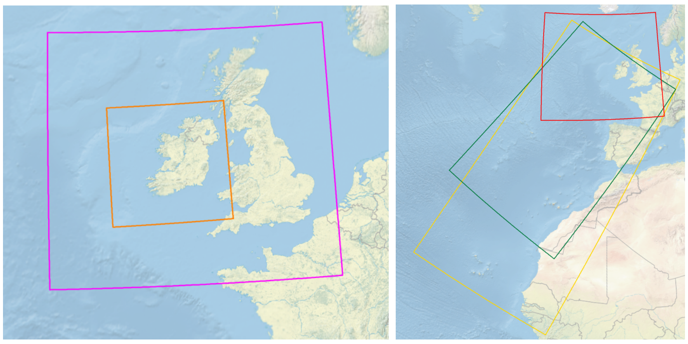

Figure 1: Domains used throughout this article. Left, for sub-kilometre testing: 750 m domain used by HECTOR (orange, 800×800) and 2 km reference for nesting (purple, 720×648). Right, Ophelia testing at 2.5 km: operational IREPS domain (red, 1000×900) along with larger 1080×2160 (yellow) and 1000×1600 (green).

2 Boundary options

The experiments carried out here all used the bugfix version of HARMONIE-AROME Cycle 43h2.2. The reference simulations use the purple domain in Fig. 1 with horizontal resolution of 2 km and the MF_90 vertical levels. In order to mimic the UWC-West configuration, a quadratic grid was used in these simulations, with a time-step of 60 s and an uncentring parameter value of XIDT=0.14. Note that experience with the UWC-West domain suggests that 50 s is a preferable time-step; however, this smaller domain here does not contain the

Table 1: Set of experimental boundary options. The last two columns show what LBC files are used in the chain for the 1200 UTC forecast of a given hectometric experiment; for example, the har3 1200 UTC uses LBC from the HARM ref0 0900 UTC, which in turn uses LBC from the IFSHRES 0000 UTC.

Name	**Host**	**LBC age**	**HARM used by 12z**	**IFSHRES used by HARM host**
ifs	IFSHRES	6-hour old	-	0600 UTC
har3	HARM	3-hour old	0900 UTC	0000 UTC
har2	HARM	2-hour old	1000 UTC	0000 UTC
har1	HARM	1-hour old	1100 UTC	0000 UTC
har0	HARM	same forecast	1200 UTC	0600 UTC
harS	HARM	same forecast	1200 UTC	1200 UTC

same complex orography, and no problems were encountered. For 3DVAR data assimilation, the UWC-West structure functions were used.

For the hectometric experiments, the 800×800 domain with 750 m horizontal resolution, shown in orange in Fig. 1, was used with the basic HECTOR set-up as follows: cubic grid truncation, MF_90 vertical levels, 30 s time-step with no uncentring, and increased spectral horizontal diffusion through RDAMPX=10 for all parameters X. Structure functions for this HECTOR domain have been generated, allowing the use of 3DVAR (this has in general led to a slight improvement in scores in the first few hours, although more neutral later on). For all of the experiments, three-hour cycling was used with a five-day warm-up period. Four times a day 36-hour forecasts were run, at the “main” hours of 0000/0600/1200/1800 UTC. Conventional observations were assimilated using upper-air 3DVAR, along with the usual surface OI.

Ideally a limited-area forecast will use LBC from the most recently available IFSHRES forecast. In an operational setting, due to the IFS schedule, the LBC will be either 6- or 9-hours old: that is, a 1200 UTC will use the 0600 UTC IFSHRES, as will the 1500 UTC. When using IFSHRES LBC directly in the hectometric experiments (as with HECTOR), this approach is also used.

When nesting a hectometric domain within a HARMONIE-AROME forecast, researchers often use the “same forecast” option, which means that the nested 1200 UTC would use the host LBC from 1200 UTC, i.e. the same forecast cycle. However, this is not practical for operations, as the nested experiment would have to wait until the host first completes. An option under our current operational set-up would be 3 hour-old boundaries; i.e. the 1200 UTC nested uses LBC from the 0900 UTC host.

As mentioned, UWC-West operations will produce an hourly forecast, thus giving more host boundary options. The suite will still use three-hourly cycling, and in order to mimic this for our tests, we define three sets of 2 km references: ref0 cycling at 0000, 0300, 0600, . . . using 6- and 9-hour old IFSHRES LBC; ref1 cycling at 0100, 0400, 0700, . . . using 7- and 10-hour old IFSHRES LBC; and ref2 cycling at 0200, 0500, 0800, . . . using 8- and 11-hour old IFSHRES LBC. Once these are available, we can then define a set of nested hectometric experiments (har3, har2, har1, har0) with reducing boundary ages. As mentioned, the “same forecast” har0 option is impractical in operations. For experimental purposes, we define a further, even less practical harS option. In this, each nesting step uses “same forecast”, so the 1200 UTC harS uses LBC from a 1200 UTC "refS" experiment, which in turn uses LBC from the 1200 UTC IFSHRES. All of these are listed in Table 1.

2.1 February 2022 test period

The main testing period used was the 10-20 February 2022, during which three named storms impacted Ireland: Dudley, Eunice, and Franklin. As we are mainly interested in high-impact extreme weather applications for very high-resolution modelling, this period is ideal. Figure 2 shows results for just the hectometric domains nested within HARMONIE-AROME experiments. Across many of the parameters we see quite a clear separation in the performance: harS is the best followed by har0 (both of which are sadly not feasible). After these, the remaining experiments with 1-, 2-, and 3-hour old LBC are quite clustered, with essentially no difference.

The differences are most notable in MSLP, and this all makes sense when we consider the chain of boundary information for each, which is outlined for the example of the 1200 UTC forecast in Table 1. The harS uses the newest IFSHRES information, followed by har0, with the rest all using much older LBC. (Note that the current operational arrival times do not allow for a later IFSHRES to be used for even the har1). It appears that ultimately (and perhaps not surprisingly) it is the age of the global model that dominates.

From Table 1 we also see that directly nesting in IFSHRES ("ifs" experiment) allows for more recent LBC to be used. This would explain the verification shown in Fig. 3. In general we see that this ifs experiment outperforms those nested in HARMONIE-AROME. The differences are smaller than reported previously (Clancy et al., 2022). However, as noted at the time the reference HARMONIE-AROME host used 65 vertical levels. All here use MF_90, which has a large impact, and thus we have a fairer comparison.

Notwithstanding the earlier discussion, we see that it is not just the global LBC age that matters. It is noteworthy that the hectometric ifs experiment still outperforms har0 in terms of temperature bias, for example.

A second period of 16-27 March 2022 was also tested. Conditions were very different: much calmer and colder, and in these conditions poor operational performance was noted in terms of night-time temperatures. However, results in these experiments (not shown) were mixed, with all boundary choices struggling equally at night-time: clearly other issues are at play.

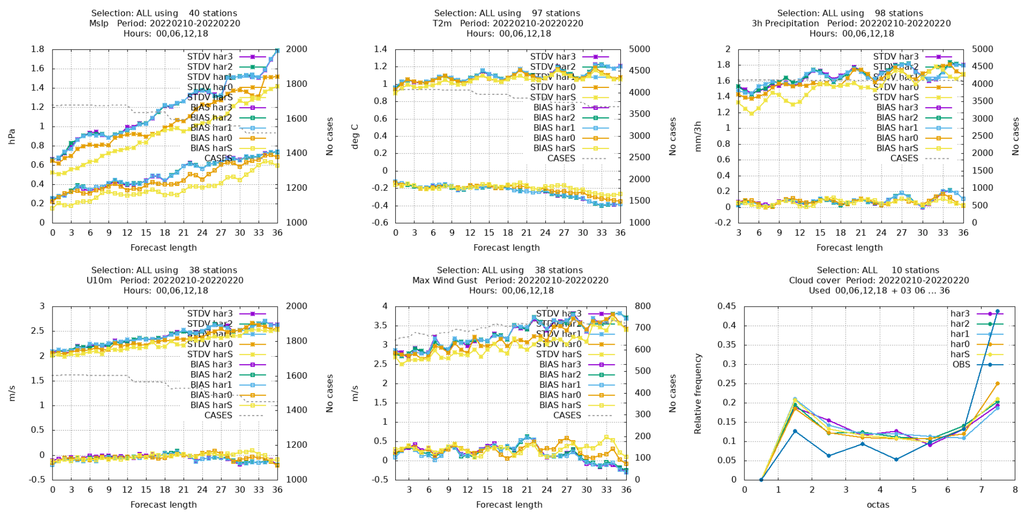

Figure 2: Verification for various parameters for the February 2022 period for hectometric experiments nested within reference HARMONIE-AROME experiments. Experiment details given in Table 1

2.2 Rainfall analysis

The previous section relied on point verification, the limitations of which are well known. Now we take a spatial view on things, with a focus on rainfall.

By default, the HARMONIE-AROME system only couples dynamical prognostic fields from the LBC files. Optional also are cloud water species: liquid (L) and ice (I); and hydrometeors: rain (R), graupel (G), snow (S). This is of huge benefit, almost essential when using the small domains required at very high-resolutions due to a dry, spin-up region near inflow boundaries (Clancy et al., 2022; Pálmason et al., 2022). Figure 4 shows plots of total accumulated rainfall over the February 2022 period for hectometric experiments with different LBC options (using the 09-09 amount from each 0000 UTC forecast). The impact of the extra coupling is quite evident at the western boundary, generally the inflow during this stormy period. Looking at the different LBC, we see that the dry boundary region remains more pronounced when IFSHRES is used (left subpanel of Fig. 4).

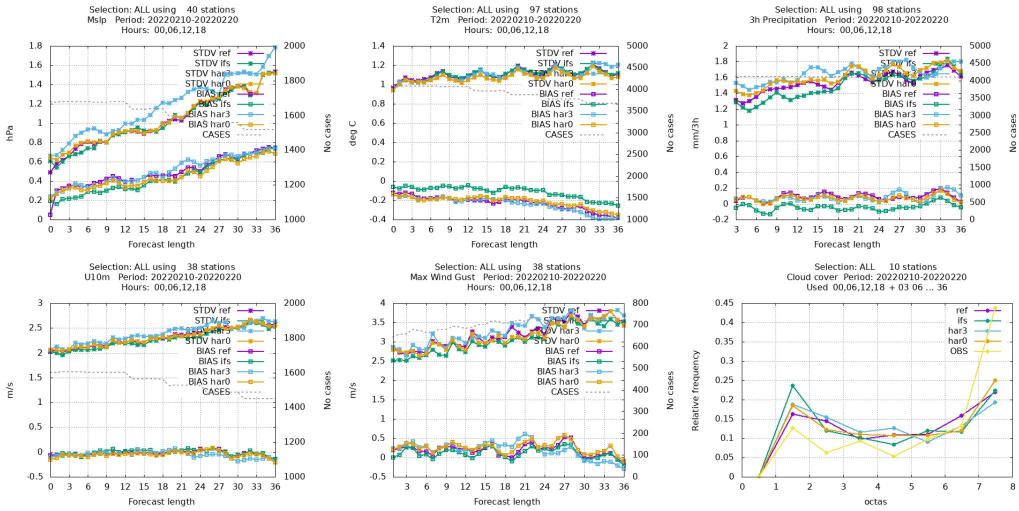

Figure 3: Verification for various parameters for the February 2022 period. Experiment details given in Table 1.

Differences are small closer to the centre of the domain, emphasising the need for as large a domain as possible. We also see that the HARMONIE-AROME driven experiments give higher totals over Ireland.

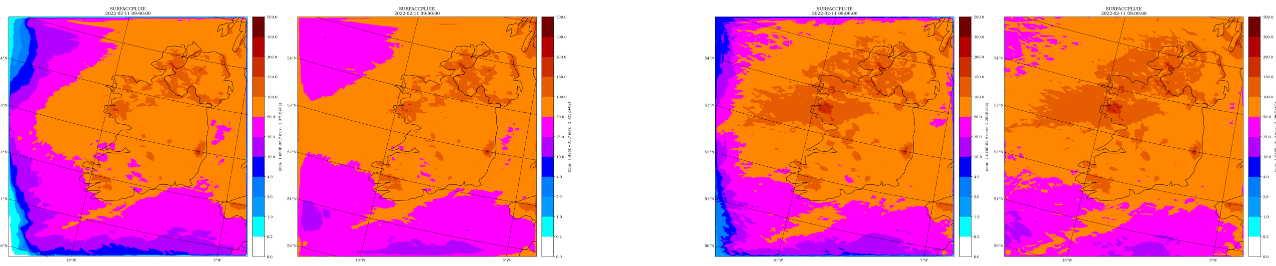

Figure 4: Total 09-09 rainfall accumulations for 10-20 February 2022, for hectometric experiment with IF-SHRES LBC (left) and HARMONIE-AROME, har3 option (right). For each is shown coupled with just cloud liquid and water (to the left), plus with all available hydrometeors (right).

We use the Irish rainfall network to evaluate accuracy, which consists of around 400 observations of 24-hour accumulations recorded daily at 0900 UTC. A gridded product at 1 km resolution is produced from this: totals for our 10-20 February 2022 period are shown on the left of Fig. 5. We see a typical Irish pattern with highest amounts in the west, particularly over mountainous regions, and also over the Wicklow mountains in the east.

Figure 6 shows the forecasted accumulations, to be compared with these observed values, from the various experiments. We would hope the higher-resolution simulations to better capture the highest amounts. In the west, this is indeed the case for the hectometric simulations using HARMONIE-AROME boundaries. The IFSHRES-driven experiment (middle panel of top row) is less successful: it is in general too dry in the west. As we saw above, less rainfall seems to be produced at the inflow boundaries with IFSHRES LBC; here the ifs experiment does a better job in the mountains on the east coast. On the other hand, we see that all experiments struggled with the amounts in the south-west.

For a more quantitative measure of accuracy, we now turn to the Fraction Skill Score (FSS), computed using the standard methodology outlined in Roberts and Lean (2008). Percentile thresholds are used in order to mitigate the impact of model biases; e.g. the 90th percentile threshold compares the location of the highest 10% of observed and forecasted 24 h rainfall. Used in this way, the FSS measures the spatial accuracy of the forecasts. Note that the entire spatial extent of the observed gridded product, as illustrated in Fig. 5, is used for the FSS

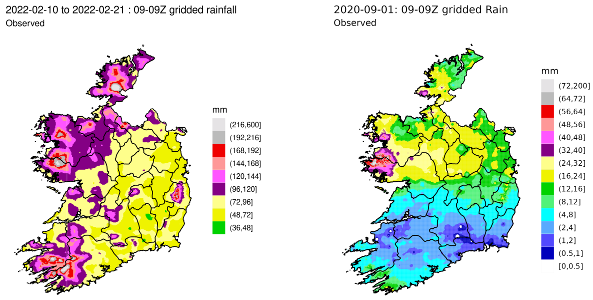

Figure 5: Gridded rainfall observations. Left: for the February 2022 period, from 0900 UTC on the 10th to 0900 UTC on the 21st. Right: for the flooding case overnight of the 1st September 2020.

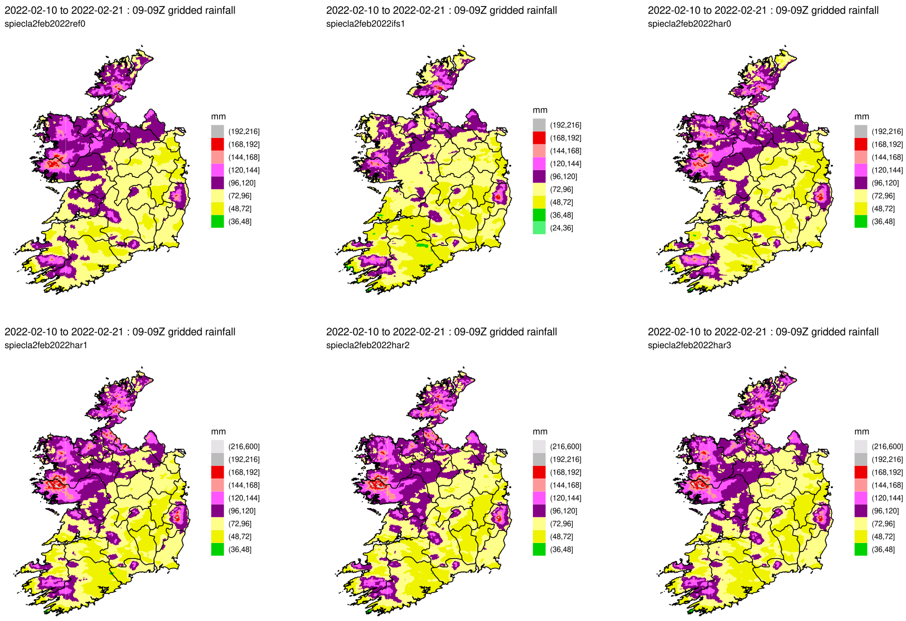

Figure 6: Forecasted rainfall accumulations for the February 2022 test period, using the 0000 UTC cycles from each day. Experiments shown are (top, left to right): ref0, ifs, har0; and (bottom, left to right) har1, har2, har3. Details in Table 1.

analysis (missing data close to the boundaries is simply omitted when upscaling).

For the February test period only forecasts from the 0000 UTC and 0600 UTC cycles at leadtimes of 33 h and 27 h, respectively, could be used for the FSS computation (giving a total of 22 samples). Results averaged over the whole period showed no significant differences. FSS performance for days with lighter (75th percentile of observed rain less than 10 mm in 24 hours) or heavier rainfall was also considered, as illustrated in Fig. 7. For lighter rainfall days (February 10, 13, 14, 16, and 18), ref0 generally outperforms the ifs and har experiments, however har1-3 tends to perform better for the heavier rainfall days (February 11, 12, 15, 17, 19, and 20). It is interesting to note that har0 is outperformed by har1-3 for the heaviest rainfall in the period (i.e. the 95th percentile on the left Fig. 7), which was also evident in standard surface skill scores (not shown). This may reflect spin-up of precipitation in the HARMONIE-AROME forecasts used for LBCs. However, one should also bear in mind the limited sample size available for this analysis.

Finally, we look at a particular event on the 1st-2nd of September 2020, when heavy rainfall overnight led to serious flooding in the Connemara region. The right-hand of Fig. 5 shows the gridded observed 24-hour

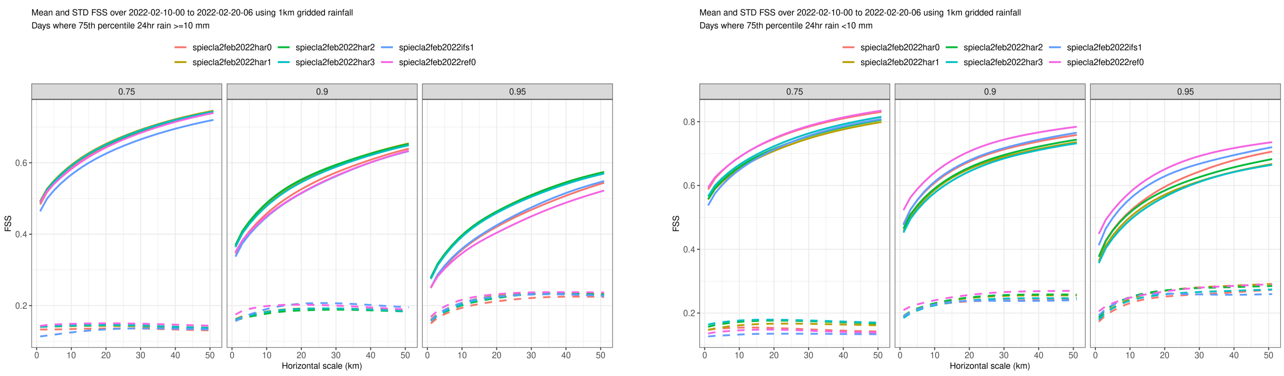

Figure 7: FSS (mean, solid, and standard deviation, dashed) for the February 2022 period, divided into "heavier" (left) and "lighter" (right) rainfall days; see text for days used. The fraction above each panel indicates the rainfall percentile. For experiment names, the last part (ifs, har0 etc) corresponds to Table 1.

accumulations valid at 0900 UTC on the 2nd. A system moved in from the west overnight, bringing highest accumulations to the mountainous areas around Connemara.

For simulations of this event the usual five-day warm-up period was used. Long forecasts from just the 0000 and 0600 UTC cycles on the 1st were used, to compare with the 09-09 observations at lead-times of 33 and 27 hours, respectively. An initial subjective view of the results showed that, once again, the age of the originating global LBC seemed to be of most importance, and so we order the results shown below in Fig. 8 by these. In this case, the "newer" global LBC actually lead to poorer performance: at the time of this event, there was considerable differences noted between runs of IFS.

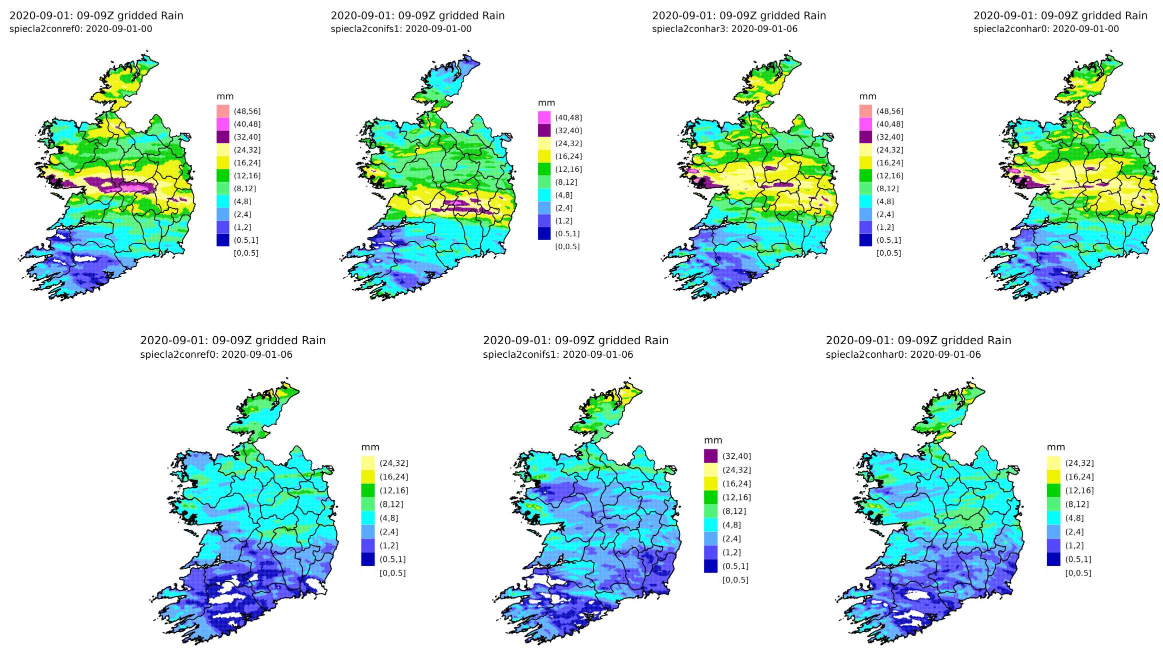

Figure 8: Experiments for the Connemara case, showing forecasted 24-hour rainfall to be compared with the right-hand panel of Fig. 5. Originating with IFSHRES LBC from 18z on the 31st (top) and 00z on the 1st (bottom). For experiment names, the last part (ifs, har0 etc) corresponds to Table 1.

Focussing on the top row of Fig. 8, we see that the HARMONIE-nested experiments (right-hand two panels) were most successful in the positioning of the heaviest rain. The IFS-driven hectometric simulation (second from left) is much too dry in the west, which is consistent with what we saw in previous sections. Given the spatial extent of this boundary spin-up we see when using IFSHRES LBC, we also extended the domain to the west to see the effect. The results were not perfect, but showed a slight improvement over the smaller domain (not shown). Increasing the number of points in the boundary coupling zone (to match the width at operational resolutions) did not significantly improve simulations either.

3 Storm Ophelia experiments

As well as the LBC tests described above, cycling experiments were carried out with other dynamics options, such as the XIDT parameter. Off-centring such as with XIDT or VESL can be used to improve stability, as does the use of a truncated grid. Averaged verification scores do not often show much effect from these options, and so a series of experiments was run to test their impact on a high-impact storm case, namely Storm Ophelia which had transitioned from a hurricane and reached Ireland on the 16th of October 2017.

The simulations again used the bugfix Cycle 43h2.2, this time on our operational domain (red in Fig. 1). The reference used all defaults: linear grid, 65 levels, 75 s time-step. No upper air assimilation was used. To see the effects at different lead-times with different boundaries, three-hour cycling was used from 2017/10/15/00 to 2017/10/16/00, with longer forecasts run at 0000 and 1200 UTC.

The changes from the default that were tested were quadratic and cubic grids, off-centring with either XIDT=0.15 or VESL=0.1, and the use of MF_90 vertical levels. We look first at point values compared with observations. The centre of Ophelia passed quite close to Valentia station in the south-west, providing a useful comparison of the intensity. Figure 9 shows forecasts from the various tests, from three different starting times. In general, we see that results were very similar, with differences due to lead-time/boundaries dominating.

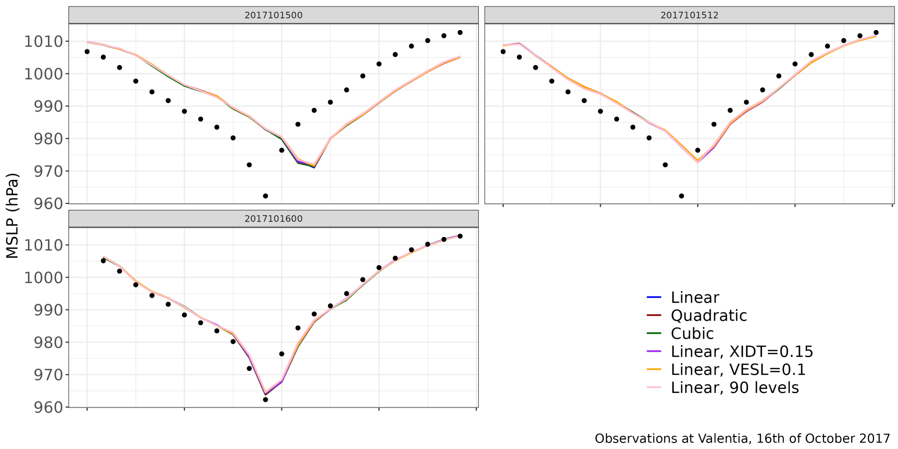

Figure 9: Forecasted MSLP on the 16th October 2017 at Valentia, along with observations (black dots). The forecast start time is given in each panel heading.

Figure 10 shows similar series for 10 m wind, but at two locations. Roches Point is on the coast and recorded the strongest winds during Storm Ophelia. Moorepark is further inland. Compared to the MSLP results, we see more differences, mainly with the use of MF_90 levels, which has more of an effect over land and reduces our peak winds in particular. This is consistent with previous findings, and has been found to be beneficial.

For a more spatial view, we consider forecasts valid at 1200 UTC on the 16th and plot the differences in fields from the default linear experiment. For MSLP, we show those a lead-time of 24 hours in Fig. 11, which includes the original fields overlaid. Differences are very small in magnitude. The MF_90 simulation shows the largest

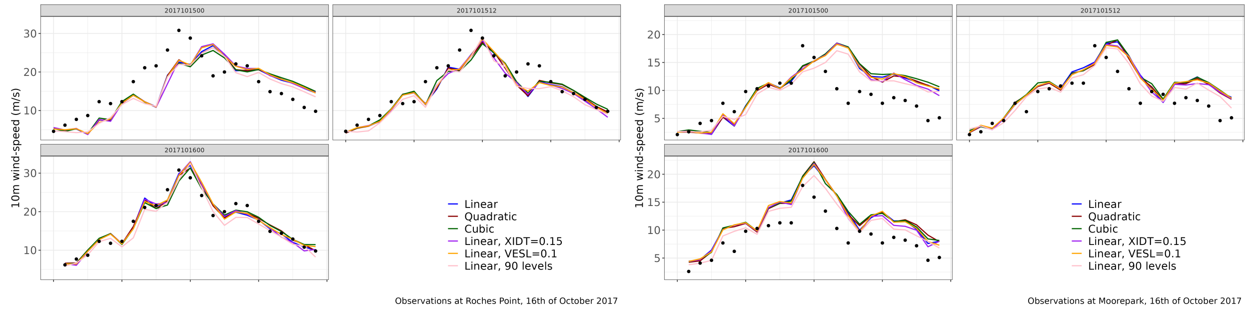

Figure 10: Forecasted 10 m wind-speed on the 16th October 2017 at Roches Point (left) and Moorepark (right), along with observations (black dots). The forecast start time is given in each panel heading.

differences, due to a very slight shift in the position of the storm.

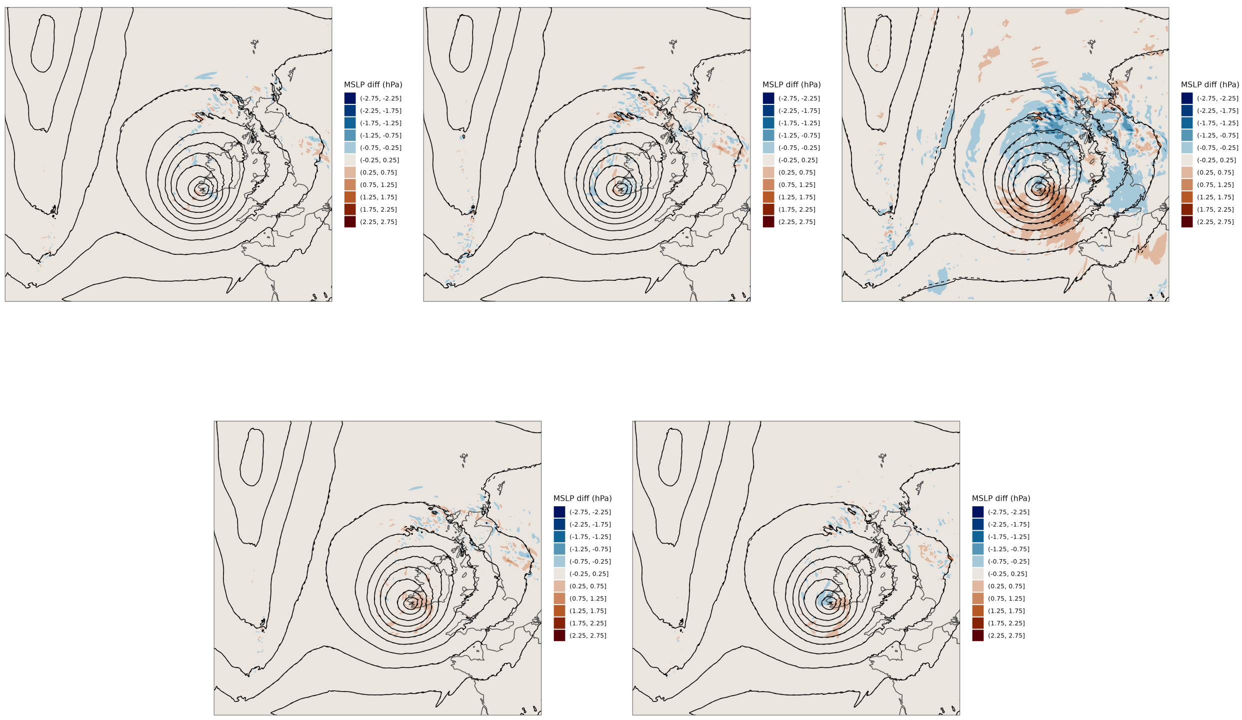

Figure 11: MSLP differences at +24 hours from the default linear experiment valid at 1200 UTC on the 16th of October 2017. For each experiment, the MSLP field is shown (solid isobars) along with the default (dashed). Top, left to right: quadratic grid, cubic grid, MF_90 levels. Bottom, left to right: VESL=0.1, XIDT=0.15

Figure 12 shows differences in 10 m wind-speed, this time for three lead-times. Most of the largest differences are due to slight positional shifts. The most obvious systematic difference is seen with the MF_90 vertical grid (middle row). As discussed above, we see reduced wind-speeds, particularly in areas where we have the strongest winds from the storm. Differences due to the off-centering options (bottom two rows) are smaller, with bigger variations due to lead-times. Longer cycling experiments did not show significant differences in scores when using XIDT, for example.

4 Summary

The main focus of these hectometric experiments was the choice of boundary conditions and the various related settings. Whether we use IFSHRES LBC directly or nest within an intermediate kilometre-scale HARMONIE-AROME experiment, it was seen that the age of the initiating global boundary files tends to dominate performance. While this is not overly surprising, it is perhaps sometimes overlooked when planning nested experiments. On the other hand, it is worth bearing in mind the practical operational constraints in terms of the delivery times of various LBC files; in general these rule out the "same forecast" nesting option.

When nesting within HARMONIE-AROME, little difference was found between 1-, 2-, and 3-hour old boundaries. In terms of point verification scores, the experiments directly using IFS LBC were found to give slightly better results. This is consistent with previous work, and is the basis for the current configuration of the HECTOR e-suite. However, previously the host HARMONIE-AROME had 65 levels. With a 90-level host, as will be available in UWC-West, the wind scores are much better and closer to those using IFS LBC.

In addition, the spatial rainfall analysis here suggests a possible benefit to nesting within HARMONIE-AROME. The simulations with IFS LBC showed a much drier spin-up region near the inflow boundaries, causing particular trouble in forecasting rainfalls in the west of Ireland. An option to mitigate this is to increase the domain size. This, of course, will come with an increased computational cost. The FSS analysis also suggested that the location of heavy rainfall may be more accurately captured when using HARMONIE-AROME-nesting. Furthermore, this nesting option is possibly wiser in the longer-term, if the resolution of the hectometric suite is to decrease further.

This work also looked at a few other dynamics settings in the case of Storm Ophelia. The use of truncated grids or uncentring did not greatly affect the intensity of the storm, and remain as attractive options for controlling the stability and cost of a forecast, particularly as resolutions increase. It is worth noting of course that these tests involved a synoptic-scale system over a domain with reasonably low-lying orography; more challenging situations would likely show more pronounced impacts from the choices.

Finally, we note some preliminary tests carried out with simulating the earlier track and development of Storm Ophela, including its post-tropical transition. These were carried out on the large domains plotted in yellow and green in Fig. 1. Cycle 46 had to be used, to benefit from the recent work in speeding up PGD preparation and climate generation, although technical problems still require the chosen domain to have "enough" land points. Stable simulations were eventually completed, and will be reported on in the future.

5 Acknowledgements

Computing resources for this work were provided through two Special Projects at ECMWF.

6 References

Clancy C., J.Fannon and E.Whelan, Hectometric-Scale Experiments at Met Éireann, ACCORD Newsletter 2, 129–138 February 2022

Pálmason B., G.N.Petersen, S.Thorsteinsson and X.Yang, Hectometric experiments with the HARMONIE-AROME system at IMO. ACCORD Newsletter 2, 139–146 February 2022

Roberts, N. and H.Lean, Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. Mon. Wea. Rev., 136, 78–97, 2008