



Assessment of MPAS variable resolution simulations in the grey-zone of convection against WRF model results and observations

An MPAS feasibility study of three extreme weather events in Europe

Matthijs Kramer^{1,2} · Dominikus Heinzeller^{3,4} · Hugo Hartmann² · Wim van den Berg² · Gert-Jan Steeneveld¹

Received: 15 November 2017 / Accepted: 27 November 2018 / Published online: 7 December 2018
© The Author(s) 2018

Abstract

Regional weather forecasting models like the Weather Research and Forecasting (WRF) model allow for nested domains to save computational effort and provide detailed results for mesoscale weather phenomena. The sudden resolution change by nesting may cause artefacts in the model results. On the contrary, the novel global Model for Prediction Across Scales (MPAS) runs on Voronoi meshes that allow for smooth resolution transition towards the desired high resolution in the region of interest. This minimises the resolution-related artefacts, while still saving computational effort. We evaluate the MPAS model over Europe focussing on three mesoscale weather events: a synoptic gale over the North Sea, a föhn effect in Switzerland, and a case of organised convection with hail over the Netherlands. We use four different MPAS meshes (60 km global refined to 3 km (60–3 km), analogous 30–3 km, 15–3 km, global 3 km) and compare their results to routine observations and a WRF setup with a single domain of 3 km grid spacing. We also discuss the computational requirements for the different MPAS meshes and the operational WRF setup. In general, the MPAS 3 km and WRF model results correspond to the observations. However, a global model at 3 km resolution as a replacement for WRF is not feasible for operational use. More importantly, all variable-resolution meshes employed in this study show comparable skills in short-term forecasting within the high-resolution area at considerably lower computational costs.

Keywords MPAS · WRF · Numerical weather prediction · Grey-zone · Voronoi grid · Convection-permitting forecast · Hail · Föhn

This paper is a contribution to the special issue on Advances in Convection-Permitting Climate Modeling, consisting of papers that focus on the evaluation, climate change assessment, and feedback processes in kilometer-scale simulations and observations. The special issue is coordinated by Christopher L. Castro, Justin R. Minder, and Andreas F. Prein.

✉ Matthijs Kramer
mv.matthijs.kramer@gmail.com

- ¹ Wageningen University, Meteorology and Air Quality Section, PO box 47, 6700AA Wageningen, The Netherlands
- ² Meteogroup, Research Department, Agro Business 99-101, 6708 PV Wageningen, The Netherlands
- ³ Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Kreuzackbahnstr, 19, 82467 Garmisch-Partenkirchen, Germany
- ⁴ Institute of Geography, Augsburg University, Alter Postweg 118, 86159 Augsburg, Germany

1 Introduction

Many traditional numerical weather prediction (NWP) models, such as the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008), are Limited Area Models (LAMs), that require boundary conditions from larger scale models. These are generally provided by a global NWP model with relatively coarse grid spacing. This technique saves tremendous amounts of computational effort. However, that comes at a cost. These LAMs also use nesting, to bridge the gap from the coarse resolution of the global model, to the desired fine resolution required for their use case, e.g. mesoscale weather modelling (e.g. Wang et al. 2012). This nesting allows LAMs to zoom in step by step to fine resolution, though with discontinuous grid size changes.

This nesting comes with certain disadvantages however. In the smaller-scale domains, sub grid scale processes are

not explicitly resolved, but parameterised, and with the discontinuous grid size changes come parameterisation differences between domains, which is unnatural. The regional model domain size must be adequate to allow spatial spin-up of small-scale features from the lateral boundary conditions (Leduc and Laprise 2009; Leduc et al. 2011; Laprise et al. 2012; Steeneveld et al. 2015). Furthermore, the parameterisation settings can differ between nested domains. Hence, the smaller domain can forecast a solution that does not match the solution of its surroundings, as determined by a larger domain. E.g., using WRF for a case study, Bukovsky and Karoly (2009) found that an inner domain shows up to 6 mm/day precipitation difference compared to its larger parent domain, over a period of 4 months (over 700 mm cumulative difference). This difference might be a result of numerical inconsistencies or of the parameterisation settings which are not scale adaptive. Obviously, the ratio between the resolved and parameterised contributions of the physical processes varies between domains. This may lead to overestimated precipitation intensity per event in domains with parameterised convection (Gadian et al. 2018). E.g. in various studies deep convection is not parameterised at grid spacing smaller than 3 km (e.g. Pennelly et al. 2014; Pilon et al. 2016). A mother domain with a coarser resolution would then use a convection parameterisation, while an inner domain with finer horizontal resolution would not. In addition, numerous studies have tested regional models using various parameterisation setups (e.g. Jankov et al. 2005; Schwartz et al. 2010; Flaounas et al. 2011; Klein et al. 2015). They found that no single combination of parameterisation works best in all cases, but large differences in precipitation rate and onset may occur.

An alternative approach is using variable grid resolution models combined with scale-aware parameterisations. The Model for Prediction Across Scales (MPAS) uses Voronoi tessellations to create irregular multigonal grid cells around grid points, to create a global irregular grid (Skamarock et al. 2012). This allows smooth transitions from coarse to fine resolution, in contrast to nesting techniques of traditional regional models. Moreover, MPAS does not need the usual grid transformation around the poles, which results in improved computational performance for global simulations compared to classic methods with a latitude-longitude grid with polar filters or spectral methods (Skamarock et al. 2012). This also requires certain parameterisation settings to be scale-aware, in particular for cumulus parameterisations. In the version of MPAS used here, this is only implemented in the Grell-Freitas (GF) cumulus parameterisation. Scale-awareness for GF means first and foremost that the tendencies rapidly decrease as resolution increases. This process is described in Grell and Freitas (2013), and applied in MPAS by Fowler et al. (2016).

Since MPAS is a global model, there are some trade-offs with respect to LAMs. A global model with the same resolution over an area of interest will have more grid cells than a LAM. Hence, MPAS involves massive computation time and output compared to regional models with similar settings. However, MPAS has proven to be efficient on supercomputers, and as supercomputers are expected to use even more cores per node in the future, this efficiency will likely increase (Heinzeller et al. 2016), making MPAS a viable option for operational use on high-performance computing clusters. Also, it is worth mentioning that because of the approximately equal area grids, the issue of reduced timestep to satisfy the computational fluid dynamics' condition for small grids near the pole, as in latitude/longitude grid is not required as in WRF. Another important aspect of MPAS is its promise for improved weather modelling. Currently, the MPAS model has mostly been optimised and extensively tested for tropical cyclones, regarding both meteorological output and computing efficiency, which incorporates modelling techniques and parameterisation schemes from the WRF model.

The current study aims to investigate the ability of the atmospheric core of the MPAS model (MPAS-A, henceforth MPAS) to properly forecast various severe weather phenomena in Europe. To our knowledge, this is the first study where MPAS is evaluated for mesoscale weather events in Europe: a synoptic gale event in north-western Europe, a föhn event in the Alps, and a case of organised convection in north-western Europe. Previous studies focussed on Northern America or the lower latitudes (Landu et al. 2014; Heinzeller et al. 2016; Pilon et al. 2016), and concluded that MPAS was capable of properly simulating the Intertropical Convergence Zone, the Western African Monsoon dynamics, and the Madden-Julian Oscillation. Also, MPAS outperforms the GFS model score for tropical hurricanes in the Pacific (Davis et al. 2016). However, MPAS has never been tested on the scale of a small European country (scale ~ 300 km). This is also the first time that real-case forecasting experiments are conducted with a global 3 km mesh with more than 65 Mio grid cells. The model runs and post-processing of four 72-h forecasts with the 3 km mesh and the other MPAS meshes described below required several million CPU hours on a large HPC system (see details in Sect. 6). The paper is organised as follows. Section 2 will give a short overview of the model settings. Sections 3, 4, and 5 evaluate the meteorological performance of MPAS for three test cases. Section 6 discusses the runtime performance of MPAS, and conclusions are drawn in Sect. 7.

2 Models, data and methods

This study evaluates WRF and MPAS for three case studies: a synoptic gale case over the North Sea, a föhn case in Switzerland, and a case of organised convection with hail in the

Netherlands. This section discusses the model setups. For spatial evaluation, we use data from the European Centre for Medium-Range Weather Forecasts (ECMWF) operational analysis at 0.125° resolution, which allows for evaluating the large-scale synoptic patterns against these pseudo-observations. Both models use the same land use map. One of the more important parameterisations is the planetary boundary layer (PBL) parameterisation. Both models use the MYNN 1.5 order TKE scheme (Nakanishi 2010).

2.1 WRF setup

This study utilizes WRF version 3.6.1, with a setup consisting of one 3 km resolution domain, centred at 2.0°E , 50.5°N . The model domain consists of 610×610 grid cells on a Lambert projection (Fig. 1). A single domain was preferred to avoid previously discussed artefacts by domain nesting, and to approach the resolution of initial conditions. WRF uses ECMWF operational forecast at $0.125^\circ \times 0.125^\circ$ horizontal resolution as initial conditions and boundary conditions. This is the resolution closest to WRF's 3 km, so artefacts due to downscaling of initial and boundary conditions are expected to be minimal. The WRF model utilises 39 vertical eta levels, concentrated in the lower troposphere

and up to ~ 30 km height. The selected parameterisations have earlier been optimised in-house for weather forecasting in Western Europe (see Table 1). In this setup, the cumulus parameterisation is disabled, as convection is supposed to be explicitly resolved at such fine resolution, although this discussion is still ongoing in literature (e.g. Gomes and Chou 2010; Krikken and Steeneveld 2012). The WSM6 microphysics parameterisation is a single-moment scheme and only predicts mass mixing ratios of the six water species (Hong and Lim 2006). In later WRF versions, a hybrid vertical coordinate system is supported, but not in version 3.6.1. As such, a terrain-following vertical coordinate system is used. Moreover, I/O quilting was used for the simulations discussed here.

2.2 MPAS setup

Here, we evaluate an unreleased version of the MPAS model (version 4.0), which includes a scale-aware GF convection scheme that becomes less active as grid cell size decreases. The model also includes the Thompson microphysics scheme that is able to cooperate with the scale-aware GF scheme. Both options have been included in the MPAS 5.0 release. Additionally, this version includes the SIONlib I/O layer for

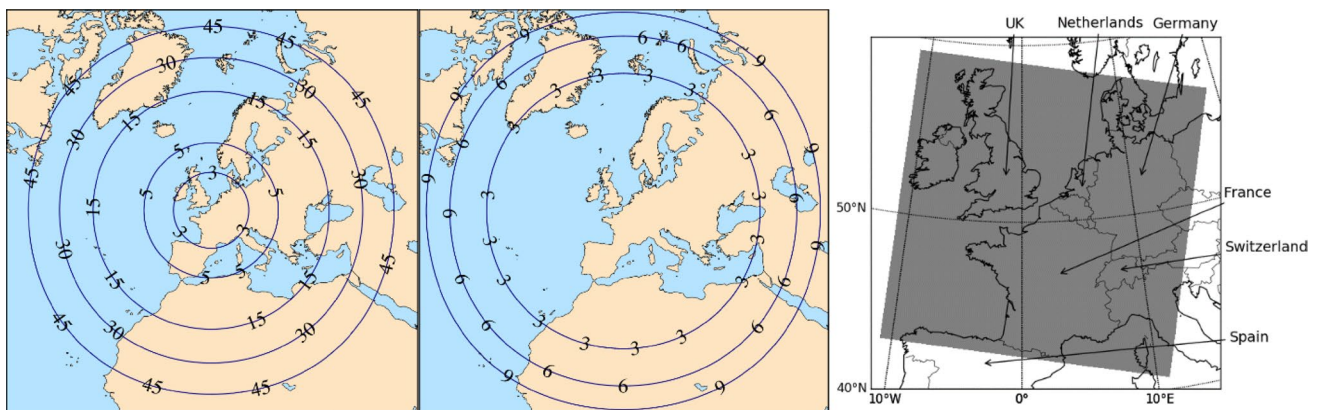


Fig. 1 Resolution refinement of the MPAS 60–3 km mesh (left), the MPAS 15–3 km mesh (middle), and the domain of the WRF model (right). The WRF model has a single domain and a 3 km resolution

throughout the domain. The refinement area for the MPAS 30–3 km is identical to the MPAS 60–3 km mesh

Table 1 Parameterisation settings for both models

Parameterisation	WRF scheme	MPAS scheme
Microphysics	WSM6 Hong and Lim (2006)	Thompson Thompson et al. (2004)
Cumulus	–	Grell-Freitas (scale-aware) Fowler et al. (2016)
Boundary layer	MYNN Nakanishi (2010)	MYNN Nakanishi (2010)
Longwave radiation	RRTMG Price et al. (2014)	RRTMG Price et al. (2014)
Shortwave radiation	RRTMG Iacono et al. (2008)	RRTMG Iacono et al. (2008)
Surface layer	MYNN Nakanishi (2010)	MYNN Nakanishi (2010)
Land-surface	NOAH Ek et al. (2003)	NOAH

improved I/O performance (Heinzeller and Duda 2017). We evaluate the MPAS model for three global grid configurations: a first grid that refines from 60 km over the globe to 3 km over Europe, a second grid that similarly refines from 15 to 3 km, but with a larger 3 km area (this grid was already used in experiments for the CONUS region), and a third grid with a uniform resolution of 3 km (Fig. 1). The global MPAS grids utilise initial conditions from ECMWF forecast data with a spatial resolution of $0.25^\circ \times 0.25^\circ$. This is an intermediate resolution between the coarsest (60 km) and finest (3 km) resolutions on the MPAS grids. The number of grid cells varies largely for the different configurations between 8.36×10^5 (60–3 km), 6.49×10^6 (15–3 km) and 65.5×10^6 (3 km). For comparison, the nested domain of the WRF setup corresponds to 3.74×10^5 grid columns.

All MPAS grids use 55 vertical levels up to ~ 30 km height, while WRF employs 39 levels up to a model top of 50 hPa. The selected parameterisations are mostly similar to WRF, though notable differences have been listed in Table 1. The utilized MPAS employs the scale-aware GF cumulus parameterisation, which means that the tendencies rapidly decrease as resolution increases. Currently, the GF scheme is the only cumulus scheme with this functionality. The Thompson microphysics scheme was selected since it is unique in its cooperation with two other parameterisations. First, Thompson microphysics has been successfully tested with the scale-aware GF scheme. Second, the Thompson microphysics scheme is uniquely coupled with the RRTMG radiation scheme concerning information about cloud droplet size distributions. We therefore consider this the most consistent approach to modelling of cloud physics. The MPAS version of the Thompson scheme used in our study is very similar to the WSM6 scheme used in the WRF setup. Both the WSM6 microphysics scheme used in the WRF model and the Thompson microphysics scheme use 6 water species. However, where WSM6 is a single-moment scheme and only determines mass mixing ratios of the precipitation species, Thompson adds another moment for rain, i.e. the number concentration, as a prognostic variable (Thompson et al. 2004). Sensitivity tests with the WRF model using the Thompson microphysics scheme provided approximately similar results as with the WSM6 scheme. Other parameterisation schemes that are identical are applied the same in MPAS as they are in WRF.

In contrast to WRF, MPAS uses a hybrid vertical coordinate system by default. While this may make the model comparison less straightforward, both vertical coordinate systems have their advantages and disadvantages. E.g. the terrain-following coordinate system can be easily coupled to the boundary and surface-layer parameterization schemes (Schär et al. 2002). On the other hand, numerical errors can occur around steep slopes (Stappeler et al. 2002), which is minimised when using a hybrid vertical coordinate system.

For each case, the MPAS model was initialised either at 00UTC or 12 UTC. For each case, the initialisation time and naming convention for the model runs will be specified. We use 15-min output intervals for the 60–3 km and 15–3 km mesh, and 30-min output for the global 3 km mesh.

3 Synoptic gale

3.1 Case description

On October 27th 2013, a depression over the North Sea with associated frontal systems caused heavy winds throughout north-western Europe. On October 28th, the cold front passed over the Netherlands accompanied by wind gusts of up to 42 ms^{-1} at the coast. Across Europe the storm resulted in power outages in several hundred thousand households and millions of euros in damages. The transportation sector was seriously affected by the extreme winds (Wadey et al. 2015; Staneva et al. 2017), and eventually 21 casualties have been reported. Obviously correct forecasting of such an event is crucial. For this case, we use five different MPAS setups and a single WRF setup. We use all MPAS grid meshes, and for the 60–3 km and 30–3 km meshes we also evaluate MPAS for two initialisation times. Model runs notations have been listed in Table 2 and will be used in the following sections of this section.

3.2 Forecast performance

3.2.1 Forecasting the synoptic situation

First, we compare forecast pressure maps on October 28 2013, 06UTC from both models to the ECMWF operational analysis, which is used as pseudo-observations since observations have been included in the data assimilation process. Next, we analyse the minimum pressure (not necessarily occurring at the same time and location) and the associated maximum wind speed.

The variable-resolution MPAS runs perform similar to the global 3 km MPAS run for the late initialisation time 28_00 (Fig. 2b–g; Table 3). In the earlier initialised run 27_12 (Fig. 2b, d; Table 3) the depression is pushed out to the east over the North Sea. This shift is even more pronounced in the 27_12 WRF run (Fig. 2h; Table 3). The 27_12 WRF run places the core more than 300 km off the ECMWF 28_00 analysis, and closest to continental Europe and the Dutch coast. The MPAS 60–3 km and 30–3 km 27_12 runs begin to shift the core westwards towards the ECMWF analysis (~ 200 km distance). All MPAS 28_00 runs place the core within less than 60 km of the analysis. The core pressure is similar for the WRF 27_12 run, the MPAS 28_00 runs and the ECMWF analysis. The earlier MPAS 27_12 runs show

Table 2 Model setup for each model run for the 2013 synoptic gale case

Model	Grid resolution				Initialisation time	
	60–3 km variable	30–3 km variable	15–3 km variable	3 km	27 Oct. 12UTC	28 Oct. 00UTC
MPAS 60–3 km 27_12	X				X	
MPAS 60–3 km 28_00	X					X
MPAS 30–3 km 27_12		X			X	
MPAS 30–3 km 28_00		X				X
MPAS 15–3 km 28_00			X			X
MPAS 3 km 28_00				X		X
WRF 3 km 27_12				X	X	

a slightly deeper core (973 hPa versus 975–977 hPa), which, in combination with the location of the core being closer to the continent, leads to slightly higher wind speeds along the Dutch coast.

The easterly shift in the low-pressure core is also present in the operational ECMWF analysis (not shown). Hence, the agreement between the model runs and the forecasts in positioning the low pressure system at 06 UTC is largely determined by the initialisation time.

The uniform MPAS 3 km run in this case does not improve over the considerably cheaper variable resolution runs. Contrary to the variable-resolution meshes, this uniform-mesh run does not employ a scale-aware cumulus parameterisation. Hence, for this test case, we conducted a sensitivity experiment with the uniform 3 km mesh in which the cumulus parameterisation was switched on. The results are almost identical to the run without the scale-aware GF scheme (not shown). This confirms that the scale-aware GF scheme is mostly inactive at a resolution of 3 km, i.e. also in the high-resolution area of the variable-resolution meshes. This is shown in more detail in Sect. 5.2.2, where both resolved and parameterised precipitation are shown separately.

3.2.2 Forecasting wind characteristics

Instead of performing a point-based comparison analysis, we opted for an extreme analysis. Here, we search for the location where the most extreme conditions, i.e. highest wind speeds and gusts, are both forecast and observed. We then compare location, and characteristics such as tendency and magnitude. For this case, we searched for the observation station in the Netherlands that recorded the highest wind gusts, which occurred at AWS Vlieland, an automated weather station owned by the Royal Netherlands Meteorological Institute (KNMI).

To identify the location of the highest wind speeds, we use a two-step approach. The highest wind speeds are frequently located to the southwest of the location and time where the depression has reached its lowest pressure

following the experience of European forecasters. Therefore, we first search for the time and location where the depression reached its lowest pressure. We limit this search to the North Sea and the Netherlands, from 53.0 to 55.0°N and 5.00–7.00°E, within a time period from 07 UTC to 10 UTC for the earlier 27_12 runs, and from 09 to 12 UTC for the other runs, including the WRF run. This difference in time window was specified to account for the easterly dislocation of the earlier MPAS runs. Subsequently, we search for the highest wind speeds around this point and time. As such, we define an area 180 km to the south and 120 km to the west of the minimum pressure location. These dimensions are chosen empirically with the area of interest (southern North Sea) in mind. Within this box, with a margin of 1.5 h around the time of minimum pressure, we search for the highest wind speeds.

All forecasts locate the low-pressure centre over the North Sea, to the north of the Netherlands (Fig. 3; Table 4). All MPAS runs are located at almost the same latitude further north than the analysis, whereas the longitude varies with initialisation time (earlier runs are further to the West) and the with the mesh employed. The WRF 27_12 run positions the lowest pressure close to the ECMWF analysis. The earlier initialised runs all forecasts the lowest pressure 2–3 h too early for both WRF and MPAS.

Interestingly, the ECMWF analysis shows the wind maximum to the east of the depression, which may be a result of the data assimilation performed by the ECMWF on the operational forecast (Fig. 3; Table 4). There is little variation in the location of maximum wind speed between the WRF and MPAS runs, as they have this location clustered. The timing and magnitude of maximum wind speeds for the MPAS runs is much improved by later initialisation towards within 1 h of the observed maximum. Of course, the comparison with the ECMWF analysis data should be handled with care, as both its spatial and temporal resolution is much coarser than any of the other model runs.

At the location of maximum wind speed obtained as described above, we compare simulated wind characteristics to observations (Fig. 4). Both models, gusts are retrieved

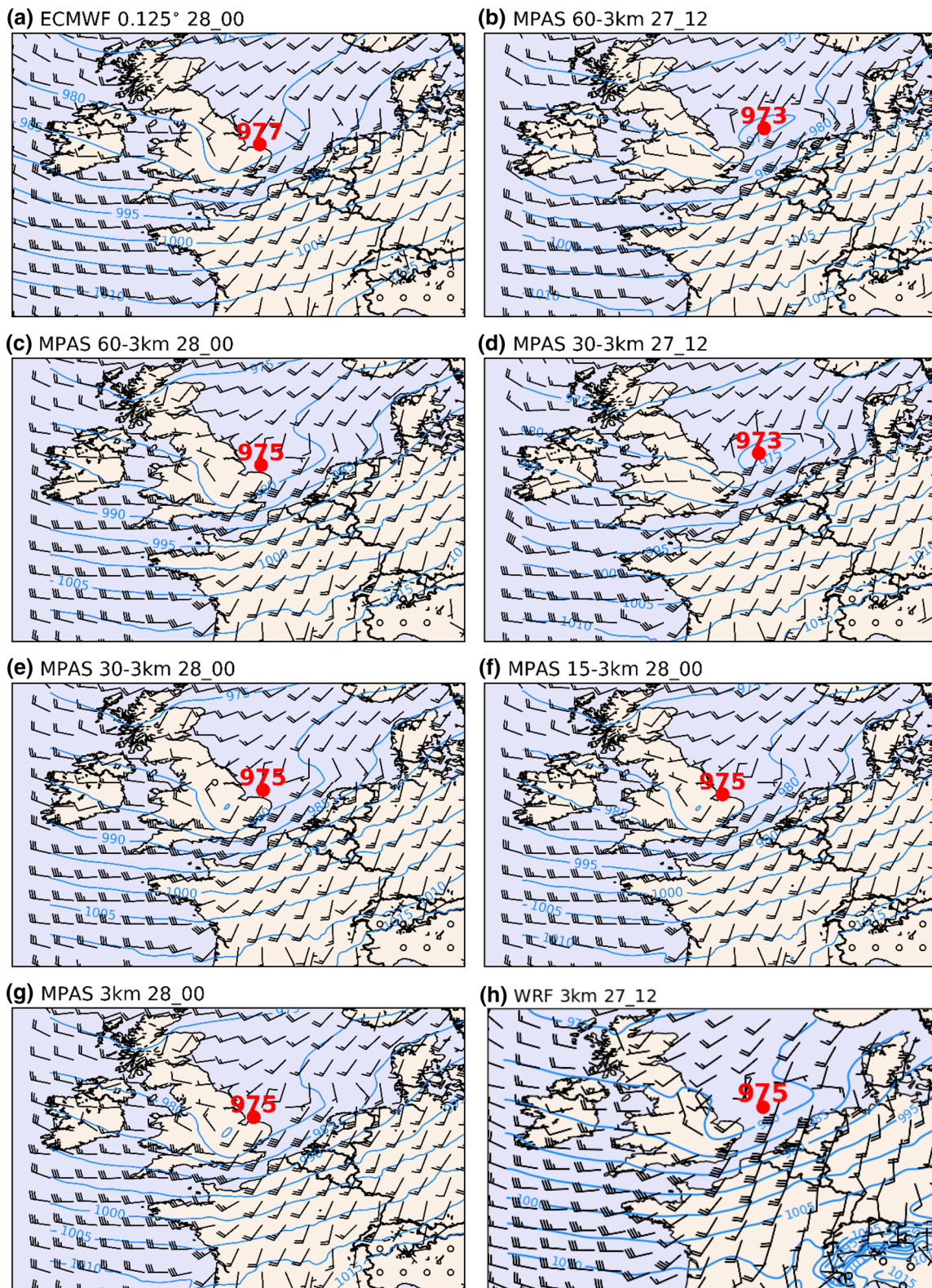


Fig. 2 Sea level pressure (blue contours) in hPa with 5 hPa contour intervals, and 10 m wind (barbs) in kts. on October 28, 2013, 06UTC. The red dot and value indicate the location and magnitude of the cen-

tre of the depression in hPa. This is not necessarily the location and time when the depression had its lowest pressure

Table 3 Location of the depression core at Nov. 28th 06UTC, for each model run, and the distance to the location of the pressure core in the ECMWF analysis data

	ECMWF	MPAS						WRF
	0.125 28_06	60–3 km 27_12	60–3 km 28_00	3–3 km 27_12	30–3 km 28_00	15–3 km 28_00	3 km 28_00	3 km 27_12
Latitude (°N)	53.13	53.97	53.34	53.97	53.34	53.13	53.24	53.75
Longitude (°E)	1.13	2.82	1.17	2.56	1.30	0.62	0.81	2.80
Distance to ECMWF (km)	–	146	23.8	133	26.6	33.5	24.6	131

Fig. 3 Modelled and observed locations of the centre of the depression (circles) and the location of the highest wind speeds per model run (triangles) in the time frame of 28 October 2013 09:00–12:00 UTC. The pressure value in hPa and the timing of the pressure minimum are noted in the legend. The black circle represents the ECMWF analysis pressure minimum, whereas the black triangle depicts the Vlieland AWS surface wind maximum (see text for details). 27_12 and 28_00 indicate model initialisation times of 27 October 2013 12:00 UTC and 28 October 2013 00:00 UTC

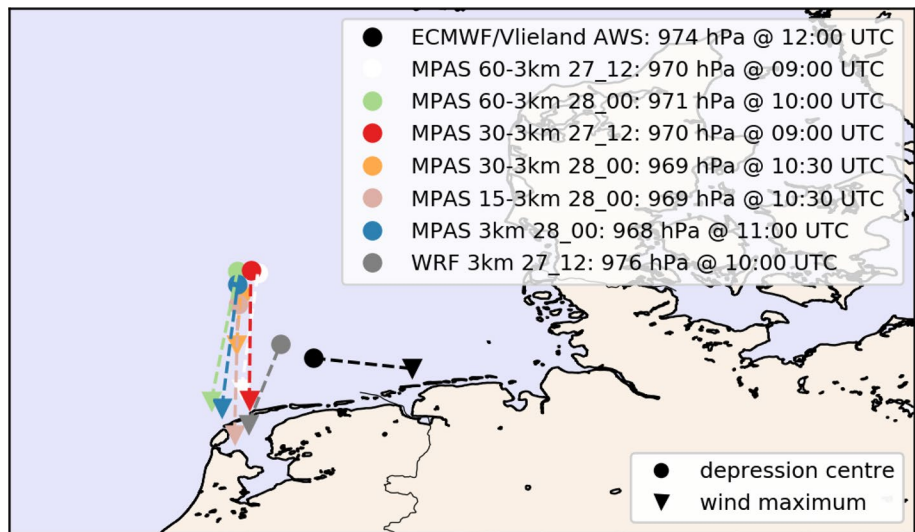


Table 4 Properties of the low pressure system when its core was at its lowest pressure, and properties of the highest wind speeds found around this point

Model run	Depression core				Wind maximum			
	Latitude (°N)	Longitude (°E)	Pressure (hPa)	Time (UTC)	Latitude (°N)	Longitude (°E)	Wind (ms ⁻¹)	Time (UTC)
ECMWF 0.125° 28_00	54.00	5.88	974.4	12:00	53.88	7.00	27.6	12:00
MPAS 60–3 km 27_12	54.98	5.24	970.5	09:00	53.64	5.05	23.3	08:45
MPAS 60–3 km 28_00	55.00	5.01	970.6	10:00	53.51	4.71	25.0	11:00
MPAS 30–3 km 27_12	55.00	5.16	970.0	09:00	53.52	5.14	22.4	08:5
MPAS 30–3 km 28_00	54.73	5.04	969.2	10:30	54.16	5.00	27.3	11:45
MPAS 15–3 km 28_00	54.61	5.02	969.4	10:30	53.13	4.98	26.7	10:15
MPAS 3 km 28_00	54.85	5.01	968.3	11:00	53.43	4.83	25.4	11:00
WRF 3 km 27_12	54.17	5.41	974.7	10:30	53.04	4.56	25.3	10:00

The automated weather station (AWS) Vlieland recorded a maximum wind speed of 35.2 ms⁻¹ (42 ms⁻¹ gusts) at 11:20 UTC

internally storing it in the next output file. For WRF however, the difference between average and maximum wind speed generally did not exceed 7%, which did not seem

realistic based on operational forecaster’s experience (not shown). Therefore, the half-hourly wind gusts from WRF are calculated according to the Beljaars algorithm (Beljaars

1987). This algorithm utilises the turbulent state of the PBL to better represent of wind characteristics. This method was unavailable for MPAS as it requires friction velocity, which is unavailable in the MPAS output. All model runs show the distinctive wind veering associated with the cold front passage, comparable to the observations. The difference between wind speed and gusts is smaller for the MPAS runs on one hand than for the WRF run and the observations on the other hand. While sustained wind speeds are higher for the MPAS 28_00 runs and closer to the observations, the wind gust calculation in WRF leads to stronger gusts more in line with the AWS Vlieland. The earlier MPAS 27_12 runs show much less pronounced wind speed patterns, but still agree well for the wind veering.

Generally, the observations show higher values for both wind speed and wind gusts (Table 5). The hourly-calculated root mean square error (RMSE) over the 24-h period is lowest for the MPAS 15–3 km run, followed closely by the global MPAS 3 km run. The time series of winds/gusts displayed in Fig. 4 shows a close match to the observations with respect to the timing of the passing of the front, the veering of the wind and the peak wind speeds. The operational

WRF 27_12 run is comparable to the variable-resolution MPAS meshes at the later initialisation 28_00. The MPAS 27_12 runs have highest RMSE values in this experiment. For MPAS, the difference between gusts and sustained winds is larger than for WRF, a consequence of the different approaches to calculate the wind gusts.

3.3 Summary

The MPAS runs show a good agreement with the ECMWF operational analysis regarding path and core pressure, comparable to the uniform MPAS 3 km run for the same initialisation 28_00. The earlier MPAS 27_12 and WRF 27_12 runs displace the low-pressure system to the east and lead to lower sustained wind speeds ahead of the recorded time. The wind gusts calculation for WRF (which could be applied to MPAS as well) improves the forecast gusts. This calculation was applied to the sustained winds from WRF, because the internally-computed gusts from the WRF modelling system showed little difference to the sustained wind speeds. In summary, the uniform MPAS 3 km mesh performs best in this test case, closely followed by the variable-resolution

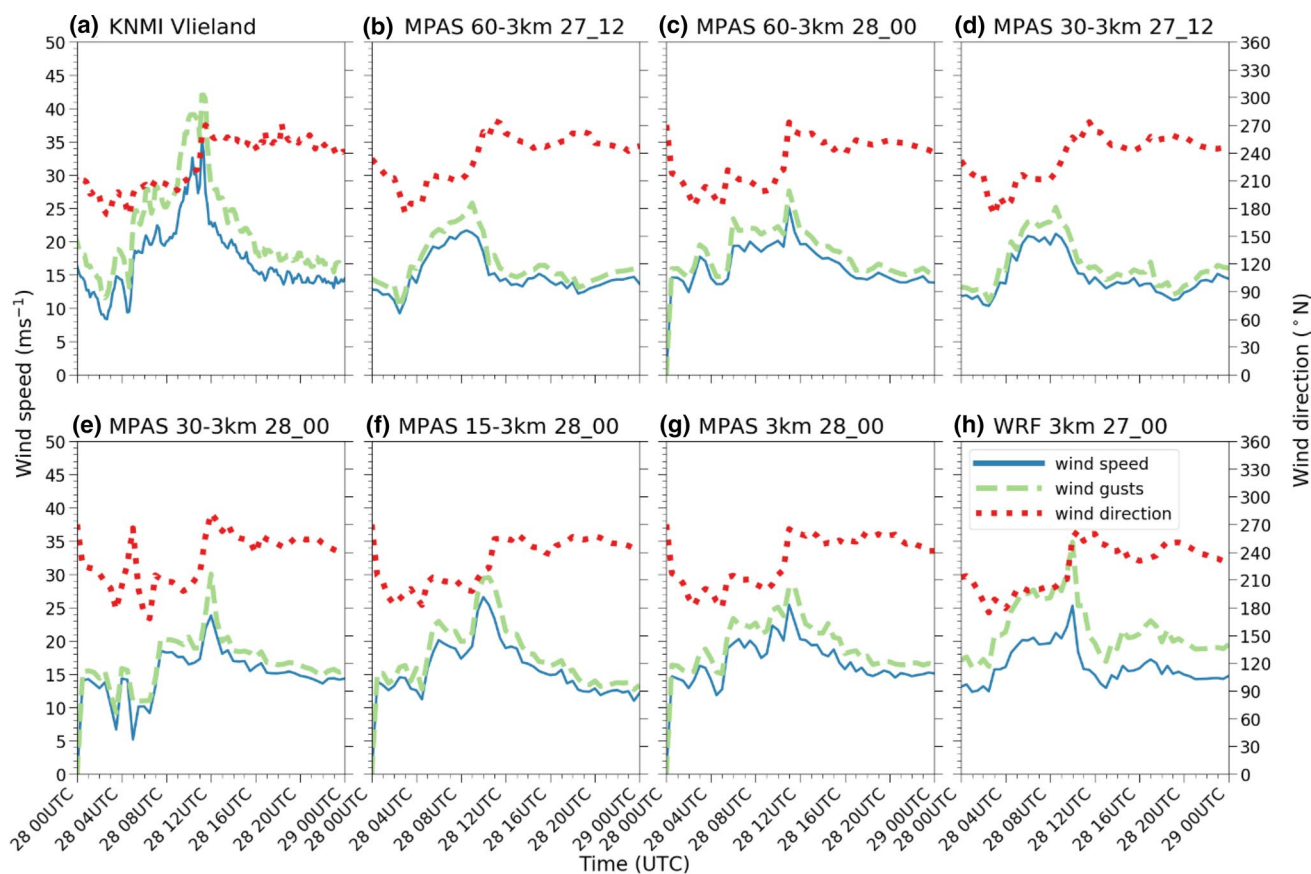


Fig. 4 Time series of wind speed, direction, and gusts per model run. The location for each graph can be found in Fig. 3. “KNMI Vlieland” is an automated weather station, controlled by the Royal Netherlands Meteorological Institute, that records data every 10 min

Table 5 Minima, maxima and root mean squared error (RMSE) of wind speed and wind gusts for all model runs and observations

Model run	Wind speed (ms ⁻¹)			Wind gusts (ms ⁻¹)		
	Min.	Max.	RMSE	Min.	Max.	RMSE
KNMI Vlieland	8.3	35.2	0	11.5	42	0
MPAS 60–3 km 27_12	9.2	21.7	4.8	10.1	25.8	7.7
MPAS 60–3 km 28_00	12.4	25	3.3	14.6	27.6	5.9
MPAS 30–3 km 27_12	10.3	21.2	5.1	11.1	25.2	8
MPAS 28_00	5.2	23.8	5.1	9.2	30	7.7
MPAS 15–3 km 28_00	11	26.6	2.6	12.7	29.5	5.2
MPAS 3 km 28_00	11.9	25.4	2.8	15.1	28.1	4.7
WRF 3 km 27_12	12.3	25.3	4.3	16.0	34.8	5.4

meshes at the same initialisation time (28_00). The earlier WRF run (27_12) and the earlier MPAS runs (27_12) suffer from a displacement of the depression to the east and to earlier hours. Nevertheless, the operational WRF run places the locations of minimum pressure and maximum wind speeds closest to the ECMWF data and to the observations. This could be related to the regular, 6-h updates of boundary conditions and sea surface temperature for WRF, while MPAS runs freely after initialisation and applies a simple diurnal cycle to the initial sea surface temperatures.

4 Föhn effect

4.1 Case description

On November 3rd, 2014, a depression over north-western Europe caused southerly winds over the Swiss Alps. This resulted in a föhn effect, where orographic lifting leads to precipitation on the windward side. The accompanied latent heat release induced a temperature difference between the lee side and windward side of the Alps. The strong winds that are associated with the föhn can be hazardous, as they arrive suddenly and with very strong gusts. The intense precipitation may trigger local flooding, and therefore timely warnings are needed with accurate forecasts. For this case, we use six different MPAS setups and a single WRF setup.

Again, we use two initialisation times for the 60–3 km and 30–3 km MPAS meshes (Table 6).

4.2 Forecast performance

4.2.1 Precipitation and temperature patterns

Since the föhn effect is caused by orography and was present for multiple days in late autumn, the accumulated precipitation may exceed several hundreds of mm, which is confirmed for all MPAS and WRF runs (Fig. 5). Patterns of simulated precipitation are very similar to each other and to observations (Fig. 5a), with the bulk of precipitation located along the Swiss-Italian border. The MPAS precipitation patterns in particular resemble each other for the same initialisation time, independent of the mesh. The earlier initialised runs (Fig. 5c, e) suggest slightly less precipitation than the remaining MPAS runs, whereas the WRF run shows the highest accumulated precipitation. All model runs clearly show locations in Switzerland that experience up to 12 °C higher night-time (06 UTC) temperatures than northern Italy, even though their elevation is several hundreds of meters higher (Fig. 6, areas with higher elevation generally have a more blue colour in these plots, or are darker green or white in Fig. 7). However, observations show that the area where the temperature is above 15 °C is higher than in any model run. Particularly for the area around Lucerne (yellow

Table 6 Model setup for each model run for the 2014 föhn case

Model	Grid resolution				Initialisation time	
	60–3 km variable	30–3 km variable	15–3 km variable	3 km	3 Nov. 00UTC	3 Nov. 12UTC
MPAS 60–3 km 3_00	X				X	
MPAS 60–3 km 3_12	X					X
MPAS 30–3 km 3_00		X			X	
MPAS 30–3 km 3_12		X				X
MPAS 15–3 km 3_12			X			X
MPAS 3 km 3_12				X		X
WRF 3 km 3_00				X	X	

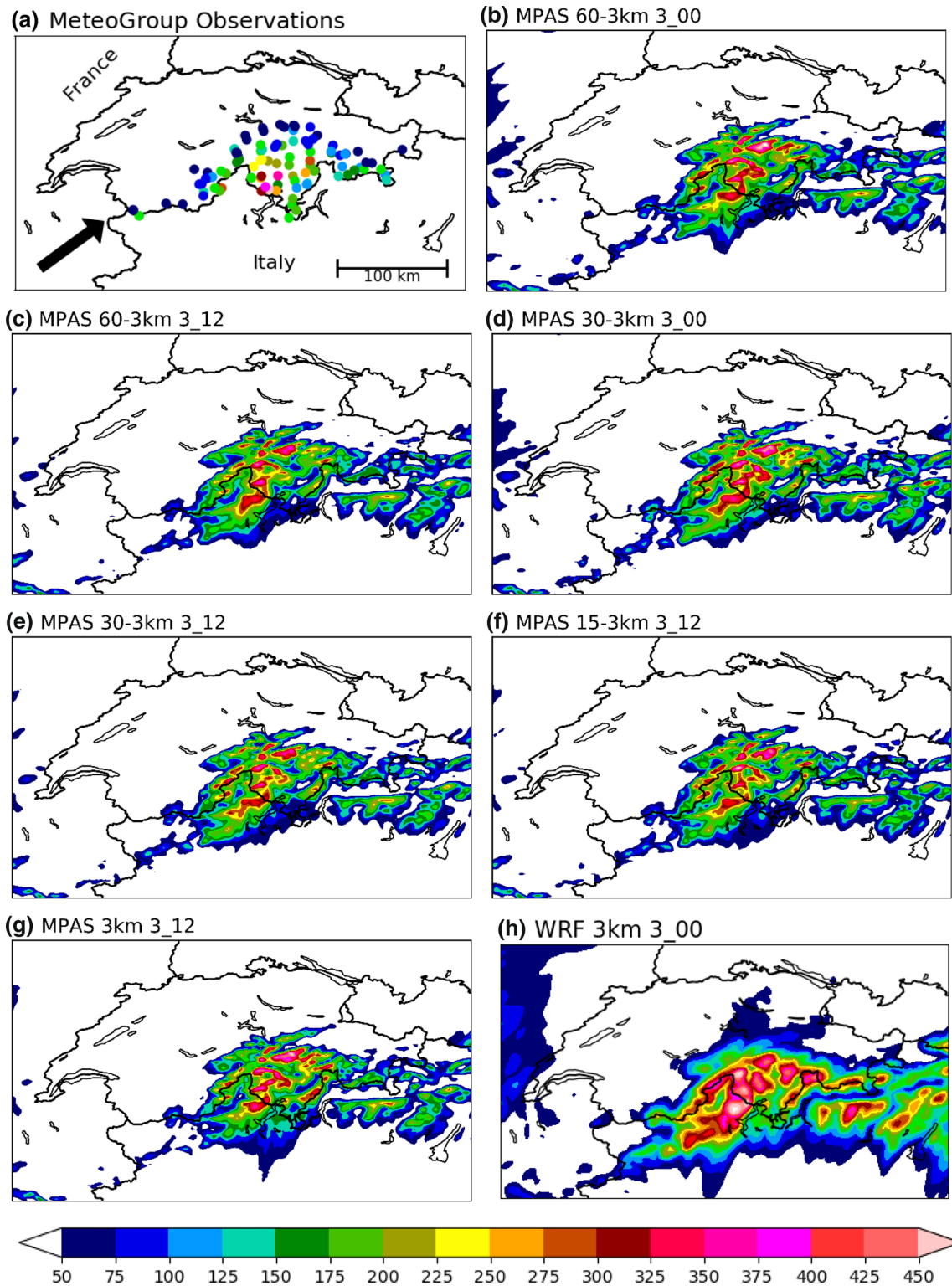


Fig. 5 Spatial maps of cumulative precipitation at Nov. 6th 12UTC for all model runs since model run initialisation. For the observations: cumulative precipitation since November 3rd 00UTC. The black

arrow shows an artist's impression of the dominant wind direction, which is valid for both the observations and all model runs

circle in Fig. 7), the WRF run and even more so the MPAS runs predict higher temperatures than the observations at Nov 4th 06 UTC.

4.2.2 Quantitative evaluation of precipitation

We expect the cumulus parameterisation to be (almost) inactive in this case, which is confirmed for nearly all simulations. Since the cumulus parameterisation was inactive in WRF, all precipitation is resolved. For all MPAS simulations, in the area of interest, non-resolved precipitation from the cumulus parameterisation accounts for no more than 5% of the total precipitation (not shown).

To quantify the temporal precipitation pattern, we search for the location where maximum precipitation was observed or modelled at the end of the föhn effect on Nov 6th 12UTC. This search is performed in the region between 44.00–48.75°N and 5.15–12.5°E (entire domain in Figs. 5 and 6). The location of maximum precipitation is almost exactly the same for all MPAS 3_12 runs, at the southern slopes of the central Swiss Alps (Fig. 7), which is at ~40 km from the observed maximum precipitation. The location of maximum precipitation in the earlier MPAS 3_00 runs is ~40 km to the Northeast, while WRF forecasts the maximum near Lake Maggiore, and closest to the observed precipitation maximum at ~20 km. The time series of modelled precipitation (Fig. 8) shows the same evolution as the WRF 3_00 and the MPAS 3_12 runs with similar accumulated precipitation between 470 mm and 520 mm, compared to the observed 500 mm of precipitation. The earlier MPAS 3_00 runs show a slightly delayed onset of the precipitation, but overall result in higher accumulated precipitation (550–630 mm).

4.2.3 Quantitative evaluation of temperature and wind characteristics

We already showed that there are locations in Switzerland where the night-time (06 UTC) temperature is higher than in northern Italy. Here, we investigate the temporal evolution of near-surface temperature and wind characteristics. We search for the location with the highest 2 m night-time (Nov. 3rd 20 UTC until Nov 4th 06 UTC) temperature in north-eastern Switzerland (46.85–47.95°N, 8.15–9.90°E; black box in Fig. 6) for all model runs (Fig. 7). For comparison, we use three WMO/MeteoGroup weather stations that measure 2-m temperature, 10-m wind speed and direction, selected by proximity to the model locations and typical characteristics of the föhn effect (see Fig. 9). These time series show that the MPAS 3_12 runs produce similar temperature and wind forecasts for the same location, irrespective of the MPAS mesh (Fig. 9b, d–f). Comparing the MPAS 3_12 runs with the earlier initialisation MPAS runs 3_00 and the WRF

3_00 run, we find that the earlier runs are more successful in forecasting the cessation of the föhn event around Nov 4th 12 UTC (Fig. 9a, c–e). This is possibly due to longer spin up or the quality of the initial field (e.g. Ducrocq et al. 2002). While WRF provides an excellent forecast for the selected station Seelisberg (WMO: 69,304), the MPAS runs tend to overestimate wind speed. Note that the accuracy of WRF or the lack thereof of MPAS for this particular station is not representative for all stations: for station Vaduz (WMO: 69,900), located roughly 100 km to the East in Liechtenstein, the föhn event prolonged and the MPAS 3_12 runs provide accurate forecasts (Fig. 9d, e).

One may argue that the stations Rueti (WMO: 69,093) and Seelisberg do not show a föhn effect, as they are located close to a lake. Both models set the lake temperature equal to the nearest sea surface temperature, as both models lack a lake parameterisation. In November, sea surface temperature can still be ~20 °C. Therefore, the models could show the effect of water temperature on surface air temperature, and not a föhn effect. However, these lakes are usually the lowest points of a valley. Therefore, the development of a very strong föhn effect is reasonable. Furthermore, the temperature time series show the distinct rise and drop, which are also more than 24 h apart. In addition, the model runs that show large amounts of precipitation also show a very consistent wind pattern during the same period that the temperature is stable. We therefore argue that the simulated phenomenon is in fact a föhn effect, and not a result of an incorrect representation of lakes.

4.3 Summary

For this case study, all model setups perform well for both precipitation patterns and cumulative amounts. However, the precipitation is delayed in the MPAS 3_00 runs, and the intensity is overestimated. Some Swiss regions reported higher temperatures than in northern Italy, even despite their relatively high altitude. Comparing grid point output within these areas with routine WMO observations shows that the temperature effect and wind direction of the föhn effect are generally well resolved, particularly for the earlier initialisation runs of WRF and MPAS. Wind speed tends to be overestimated in the MPAS runs for two of the three selected stations in the downwind area of the Swiss Alps.

5 Organised convection

5.1 Case description

On August 30th 2015, a warm front passed the Netherlands from the South-East towards the North Sea. The front involved atmospheric instability, heavy thunderstorms

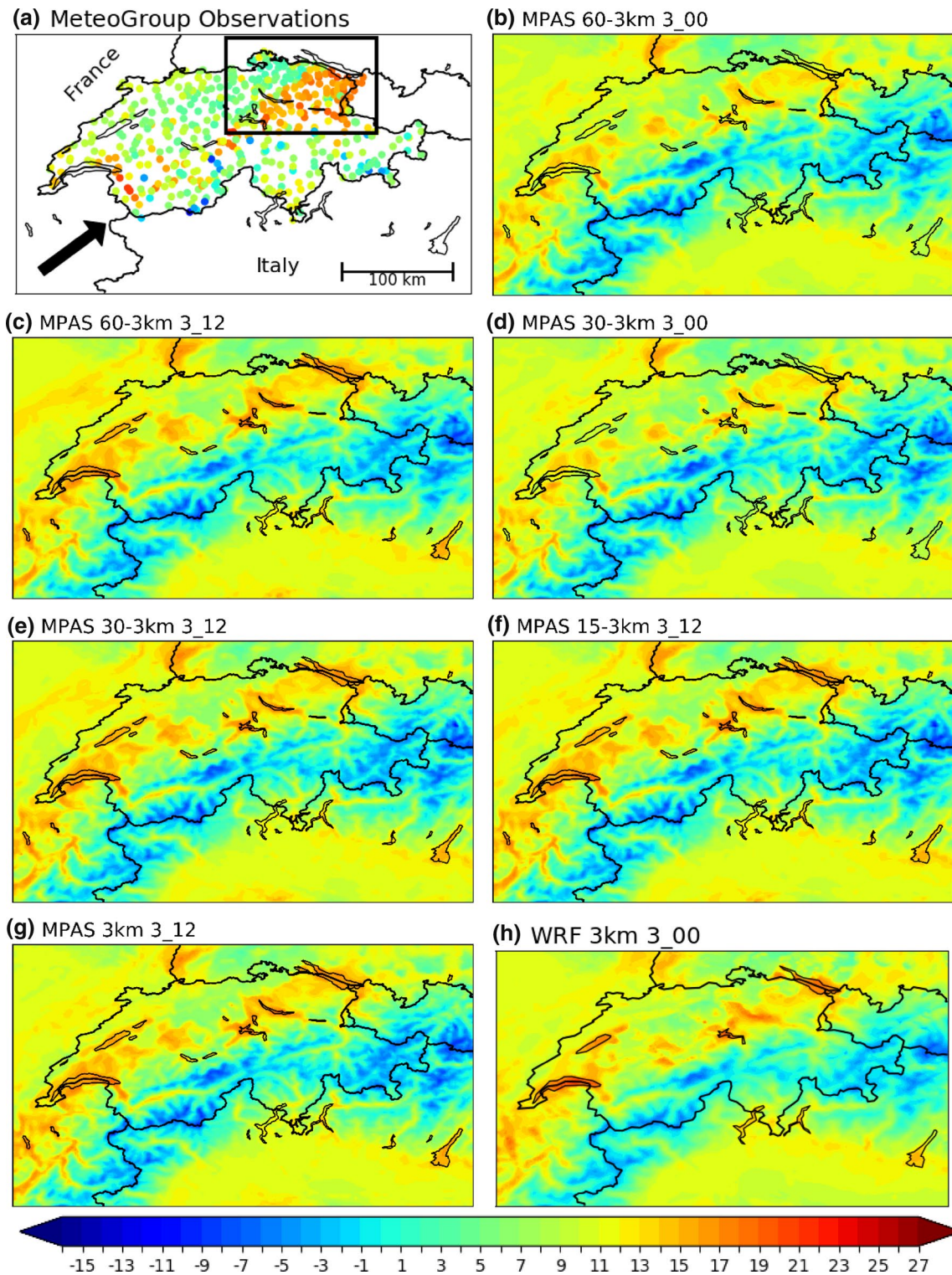


Fig. 6 Spatial maps of 2 m temperature (°C) on Nov. 4th 06UTC from MeteoGroup weather stations and for each model run. The black box in the observations window shows the area where we searched

for maximum temperatures (see text for details). The black arrow shows an artist's impression of the dominant wind direction, which is valid for both the observations and all model runs

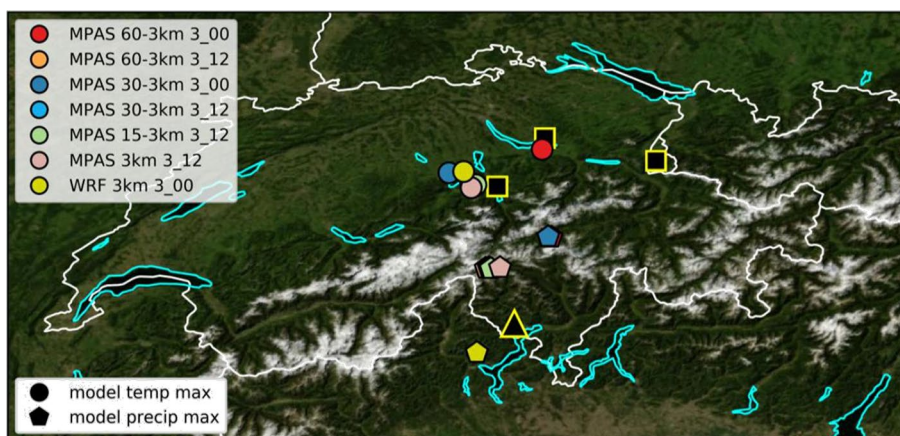
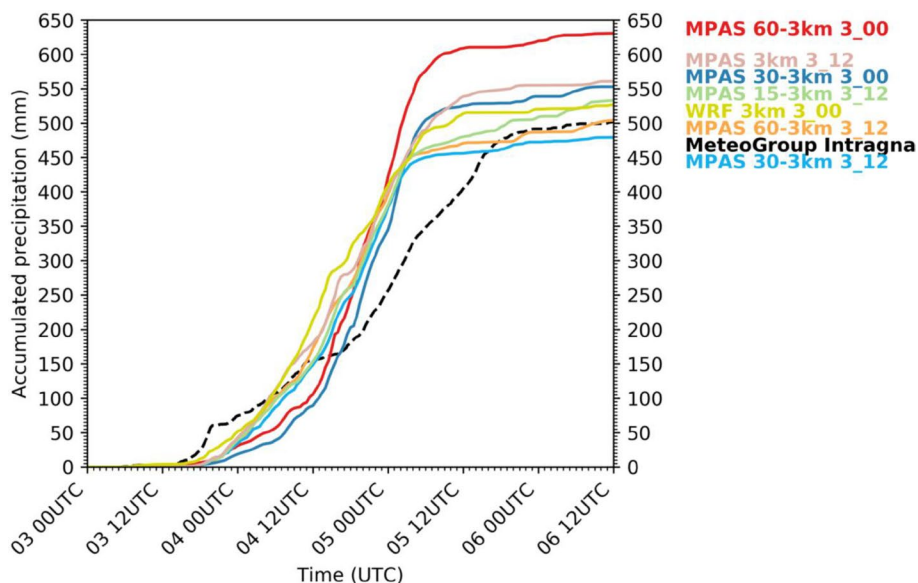


Fig. 7 Locations of maximum night-time temperature (Nov. 3rd 20:00 UTC until Nov 4th 06:00 UTC) per model run (circles), the location of maximum precipitation at Nov 6th 12UTC per model run (pentagons) and the weather station with the highest cumulative precipitation at Nov 6th 12UTC (black triangle padded yellow). Also

displayed are three weather stations used in the time series plots in Fig. 9. Note that the location of maximum night-time temperature and maximum precipitation for the different MPAS 3_12 runs are very close. Therefore, these points are overlaid on the map

Fig. 8 Cumulative precipitation time series for each model run and the observation station. Locations of the station and the point taken for each model run can be found in Fig. 7



and hail. Due to the weak upper air flow, this system was rather stationary, resulting in accumulative precipitation up to 97 mm. Also, hail stones with a diameter of 6 cm were observed, damaging mostly vehicles. In the Netherlands, $\sim 7 \times 10^4$ lightning discharges were observed (KNMI 2015). Especially for the high accumulative precipitation, a good quality forecast is crucial for timely precautions by water authorities. A forecast for intense hail storms could alert residents to avoid injuries or minimise damages to property. As hail is not a prognostic variable in versions of WRF and MPAS used in this study, we focus on the timing and amount of precipitation instead. For this case, we use the same six MPAS setups and a single WRF setup

as before. Again, we test two initialisation times for the MPAS 60–3 km and 30–3 km meshes (Table 7).

5.2 Forecast performance

NWP models generally have difficulties with properly forecasting the location and amount of convective precipitation, as it is the result of a complex feedback process between the atmosphere and the land surface (Ebert et al. 2003; Kunstmann and Stadler 2005; Case et al. 2011; Cuo et al. 2011; Eden et al. 2012; Klein et al. 2015; Gadian et al. 2018). Quantitative precipitation forecasting is therefore a much-discussed subject (Ek et al. 2003; Ebert et al. 2007; Smiatek

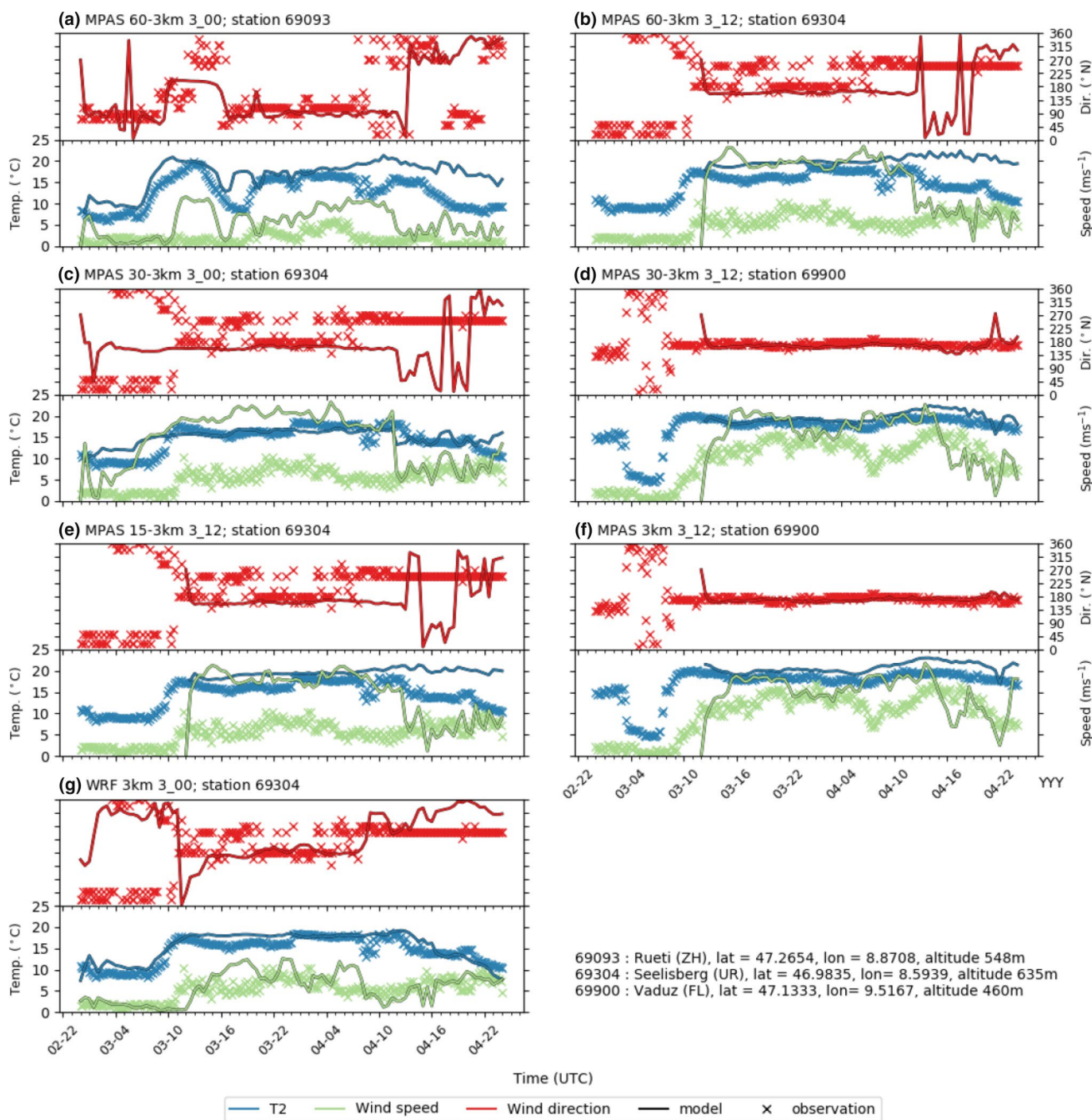


Fig. 9 Modelled 2m temperature, 10-m wind speed and wind direction time series per model run (lines), compared to observations (crosses) from a sample of weather stations with different durations of the föhn event. Tick mark units for wind speed are equal to those of temperature. Station 69,093 is close (7 km) to the location of maximum temperature of the MPAS 60–3 km 3_00 run (see Fig. 7);

station 69,304 is close to the locations of maximum temperature of WRF (17 km) and the remaining MPAS runs. Station 69,900 in Liechtenstein is 100 km to the North-East and shows a prolonged föhn period. Note that the MPAS 3_12 runs do not start at the start of these time series and that all variables require a spin up. The temperature scale is equal to the scale of the wind speed

et al. 2012). A model skill in accurately forecasting precipitation is even more uncertain at resolutions coarse enough to justify convection parameterisation (≥ 8 –10 km). In this case, the cumulus parameterisation scheme can have a tremendous effect on the patterns and amounts of precipitation

that is forecast (e.g. Crétat et al. 2012; Alam 2014; Pei et al 2014; Wootten et al. 2016). Since all MPAS runs have a resolution of 3 km over the study area, it is generally accepted that a cumulus scheme should be less active or even inactive (Hong 2004; Lim et al. 2014).

Table 7 Model setup for each model run for the 2015 organised convection case

Model	Grid resolution				Initialisation time	
	60–3 km variable	30–3 km variable	15–3 km variable	3 km	30 Aug. 00UTC	30 Aug. 12UTC
MPAS 60–3 km 30_00	X				X	
MPAS 60–3 km 30_12	X					X
MPAS 30–3 km 30_00		X			X	
MPAS 30–3 km 30_12		X				X
MPAS 15–3 km 30_12			X			X
MPAS 3 km 30_12				X		X
WRF 3 km 30_00				X	X	

5.2.1 Precipitation patterns

All model runs show precipitation patterns and amounts that correspond to observations. KNMI provides spatial maps of 24-h cumulative precipitation, measured by radar, from 8 to 8 UTC the following day. We compare this to spatial maps of 24-h cumulative precipitation per model run (Fig. 10). The 24-h window for the model precipitation is chosen so that the bulk of precipitation is located within this window. For all models, this window starts at October 31st 00UTC and ends at November 1st 00UTC. A direct comparison of model precipitation patterns with the KNMI radar image is hampered by the different time periods, in particular over the North-West of the area displayed in Fig. 10, but still provides useful guidance of the location and amount of maximum precipitation. The earlier MPAS runs and the WRF run, initialised 30_00 show a more widespread area of high precipitation amounts with WRF forecasting most precipitation, though most over the North Sea and south-eastern UK, rather than over the Netherlands. Even though the precipitation amounts agree with observations, one may argue that this is a result of the fact that the frontal system was enhanced with more water vapour from the North Sea, as sea surface temperature was between 17.6 and 19.8 °C late August 2015 (KNMI 2017). The later MPAS 30_12 runs forecast a narrow band of intensive precipitation running from South-West to North-East that coincides nicely with the radar image (note the rotation of the radar image). The total precipitation of the MPAS 30_12 runs within that band is less than in the radar observations, but more precipitation is forecast over the Netherlands than for the 30_00 WRF and MPAS runs. These 30_12 runs also show little to no precipitation further inland, in agreement with the radar observations.

5.2.2 Quantitative evaluation

As this was a convective event, a simple time series at one fixed location for all runs does not represent the properties of the event. Instead, we search for the point

of highest precipitation between 50.5–55.0°N and 2.50–8.00°E (roughly covering The Netherlands), and calculate the 95-percentile of precipitation of a square box around this point of 30 × 30 km size (around 0.35° × 0.35°) per time step. We compare this to the weather station in the Netherlands that recorded the highest amount of cumulative precipitation from August 30th 00UTC to September 1st 00UTC: KNMI Herwijnen AWS (Fig. 11). The observations show light precipitation in the beginning and heavy precipitation between August 30th 18UTC and midnight. The early light precipitation is reproduced by the WRF and MPAS runs started at 30_00, although WRF overestimates the accumulated precipitation by about 15 mm. Concerning the second event that starts between 30_18 and 31_00 UTC, the WRF and the MPAS 60–3 km runs provide a rather good forecast of the accumulated precipitation. The MPAS 30–3 km run slightly underestimates the accumulated precipitation of this event (~ 10 mm). Moreover, we find that the WRF simulation provides a much higher precipitation intensity at the beginning of the event than MPAS. This is opposite at the end of the event, where MPAS simulations show a sudden end and WRF shows a gradual end of the precipitation. However, the locations of maximum precipitation in these cases are over the North Sea and ~ 350 km from the AWS (Fig. 11). The later MPAS 3_12 runs also show this sharp rise in convective precipitation, but delayed by 6 h (60–3 km/30–3 km) to 8 h (15–3 km/3 km). The maximum precipitation is forecast over land (except MPAS 60–3 km) and almost identical to the observed value. Interestingly, the point of maximum precipitation of the MPAS 3 km run is located on the northern Dutch coast, far away from its position in other model runs and from Herwijnen. However, this point is an isolated outlier, and there is another point that has only 5 mm less precipitation at 4.912°E 52.64°N, much closer to the point of the other model runs. A similar situation is found for the MPAS 15–3 km run. The enhanced delay of the convective precipitation event in the higher-resolution MPAS runs 15–3 km/3 km can be explained as follows: for the 60–3 km and 30–3 km runs, the high-resolution

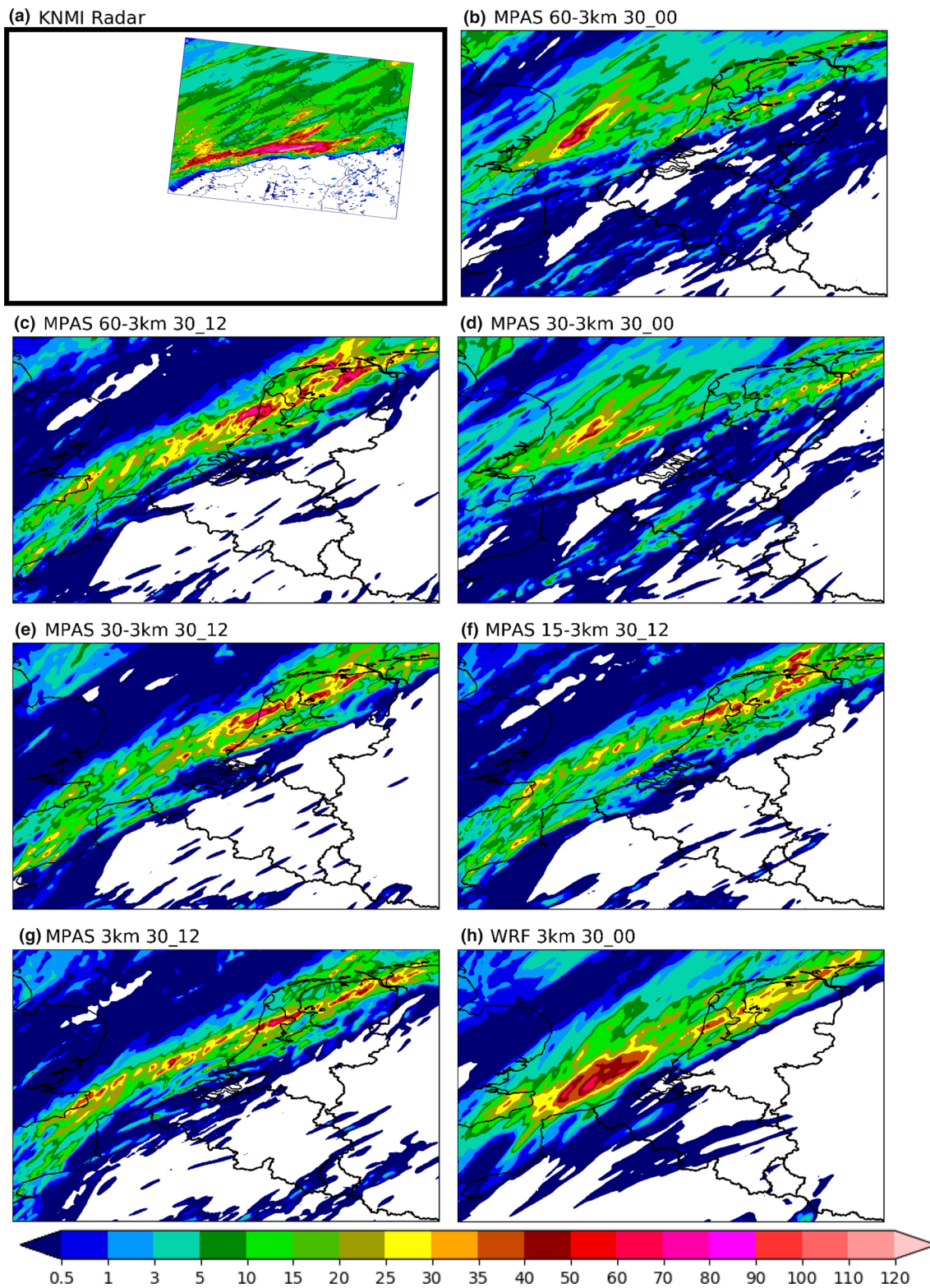
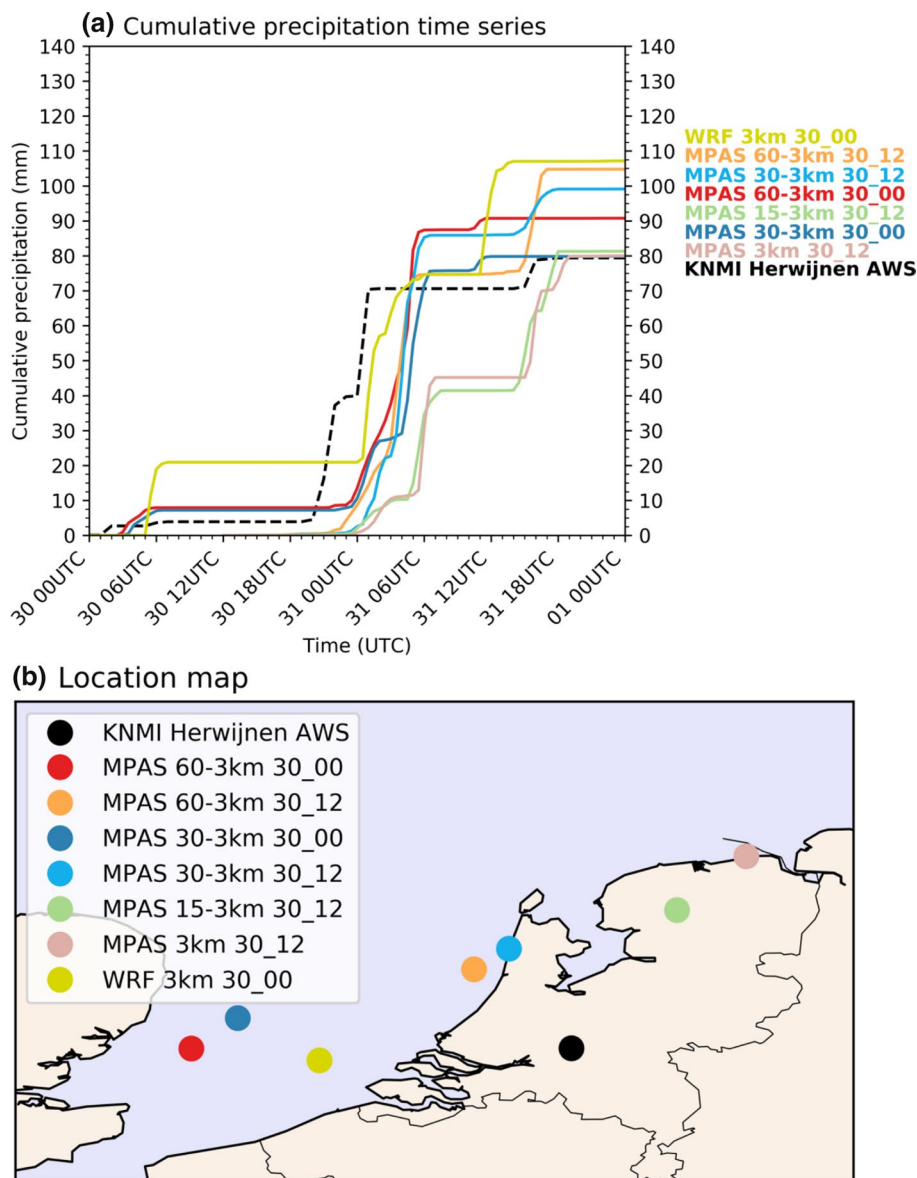


Fig. 10 Spatial maps of 24 h cumulative precipitation (mm). For the WRF and MPAS model runs, the accumulated precipitation is obtained for the period August 30th 00UTC to 31st 00UTC. The KNMI Radar images are provided daily for periods from 08UTC

the previous day to 08 UTC, here the accumulated precipitation is shown from August 30th to 31st 08UTC. The black box represents where this radar image lies relative to the other plot windows, and is included to make the visual comparison easier

Fig. 11 Cumulative precipitation time series **a** of the weather station Vlieland AWS and all model runs, and the location of the points used for the time series **b**



domain in which the scale-aware GF scheme is inactive (see below) is relatively small. Within 12 h since model initialisation, air masses from coarser-resolution areas, in which the GF scheme is active, can transit into the high-resolution area. It generally takes longer to reach saturation and the triggering of resolved precipitation when not using cumulus parameterization. Consequently, more moisture is contained in the air masses moving over the area of investigation in the 60–3 km and 30–3 km meshes and trigger the explicit convective event earlier than for the higher resolution MPAS runs.

Within the high-resolution area, the scale-aware GF scheme is nearly inactive and produces only negligible amounts of precipitation (Fig. 12). As all model runs reached Convective Available Potential Energy values of

$> 3500 \text{ Jkg}^{-1}$, we can conclude that the event was indeed convective in all model runs.

5.3 Summary

All WRF and MPAS runs performed well in this case. The later MPAS 30_12 runs show the distinctive pattern of convective precipitation within a narrow band of high precipitation that is similar to observations. The earlier WRF and MPAS 30_00 runs spread the precipitation over a larger area and have their highest amounts of rainfall over the North Sea instead of over the Netherlands. Contrary to the previous cases, the coarser-resolution MPAS meshes show a better timing of the convective event than the high-resolution meshes. This may be related to the size of the

high-resolution area and the different times required to spin up the atmospheric state with or without convective parameterisation.

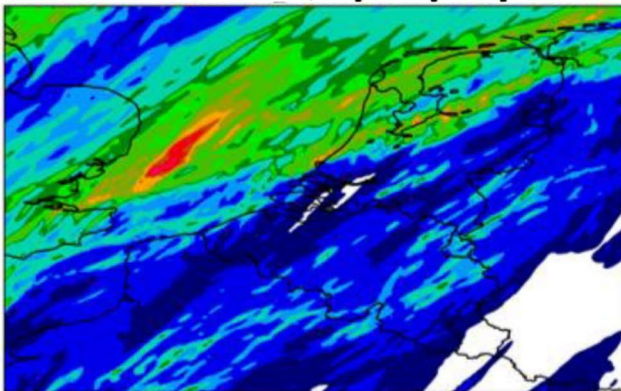
6 Runtime performance

Besides the ability to provide accurate forecasts, a key element in real-time NWP is the time to delivery and the required computational resources. For atmospheric models such as WRF or MPAS, the lower limit of nodes required to generate a forecast is constrained by the available memory to hold the large number of fields for a given problem size. For both models, the pre- and post397 processing steps generally have lower memory requirements and the runtimes are short, compared to the actual model run. We therefore focus on the model run only in the following. For reference, the model chain for a complete MPAS forecast is shown in Fig. 13. An MPAS model run consists of a model setup phase and a time integration. The costs associated with the model setup are fixed and amortised for typical integration times in weather forecasting. In recent studies, Heinzeller et al. (2016) and Heinzeller and Duda (2017) have shown that the real-time performance of the time integration (i.e. the dynamical solver, the physics parameterisations and the file output) is mainly a function of the number of grid columns per (MPI) task (the number of grid cells for which a core is “responsible”) for a wide range of parallelisations. When the number of grid columns per task falls below about 150–200, the parallel performance of the time integration decreases. It should be noted here that this relation holds for all the meshes employed in this study, as long as a file format suitable for massively parallel applications is chosen (SIONlib in this case, see Heinzeller and Duda 2017 for details). For the various netCDF file formats known to MPAS (serial netCDF, parallel netCDF4 via HDF5, and PnetCDF), the I/O

performance does not scale with the number of MPI tasks and therefore becomes a bottle neck for large meshes (Heinzeller et al. 2016). The capability to read and write SIONlib files was added to the version of MPAS used in this study.

Table 8 summarises the runtimes for the different MPAS meshes and setups used in this study, and provides scaling results for the time integration using the 60–3 km mesh. These results show that the scaling of the model is a function of grid columns per task and independent of the mesh. Figure 14 displays the scaling of the time integration (dynamics, physics, file output) for the 60–3 km mesh on the Leibniz Centre for Computing (LRZ) SuperMUC. For more than approximately 200 cells per task, the scaling is close to ideal. For larger numbers of tasks (smaller numbers of cells per task), the MPI communication overhead limits the scalability of the model. Based on the scaling behaviour of the time integration, one can calculate the required computational resources under the constraint of a fixed time to delivery. We refer to the operational WRF setup at MeteoGroup, which produces a 36-h forecast in 4 h with a total of 70 cores (MPI tasks). Table 9 lists the configurations required to create a 36 h-forecast for the three meshes on SuperMUC, assuming a similar time to delivery than the operational WRF setup. Under these constraints, delivering a 36-h forecast in 4 h requires at least 3927 nodes (62,832 tasks) on SuperMUC. While this is an impressive number, it is within reach of next generation HPC systems that will be available to the largest operational weather services such as the ECMWF. Global, convection-resolving atmospheric simulations can therefore be expected in the next five or so years to come, provided a model with good parallel scaling like MPAS is used. While such tremendous computational facilities will be available to a small audience, most operational centres and end users of MPAS will have to revert to one of the variable-resolution meshes employed here. The 15–3 km mesh requires less than 10% of the computational resources compared to the global

(a) MPAS 60-3km 30_00, explicit precipitation



(b) MPAS 60-3km 30_00, param. precipitation

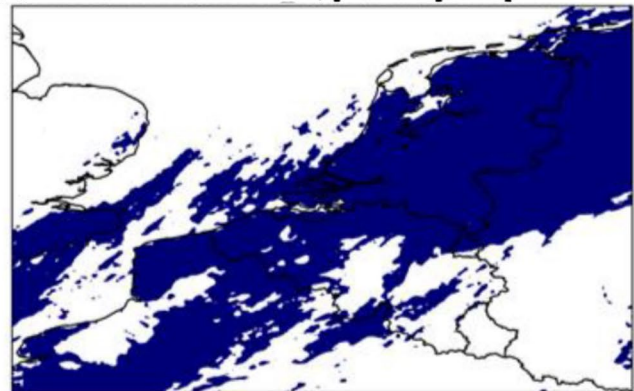


Fig. 12 Explicit versus parameterised precipitation for the MPAS 60–3 km 30_00 run, accumulated over a 24-h period from August 31st 00UTC

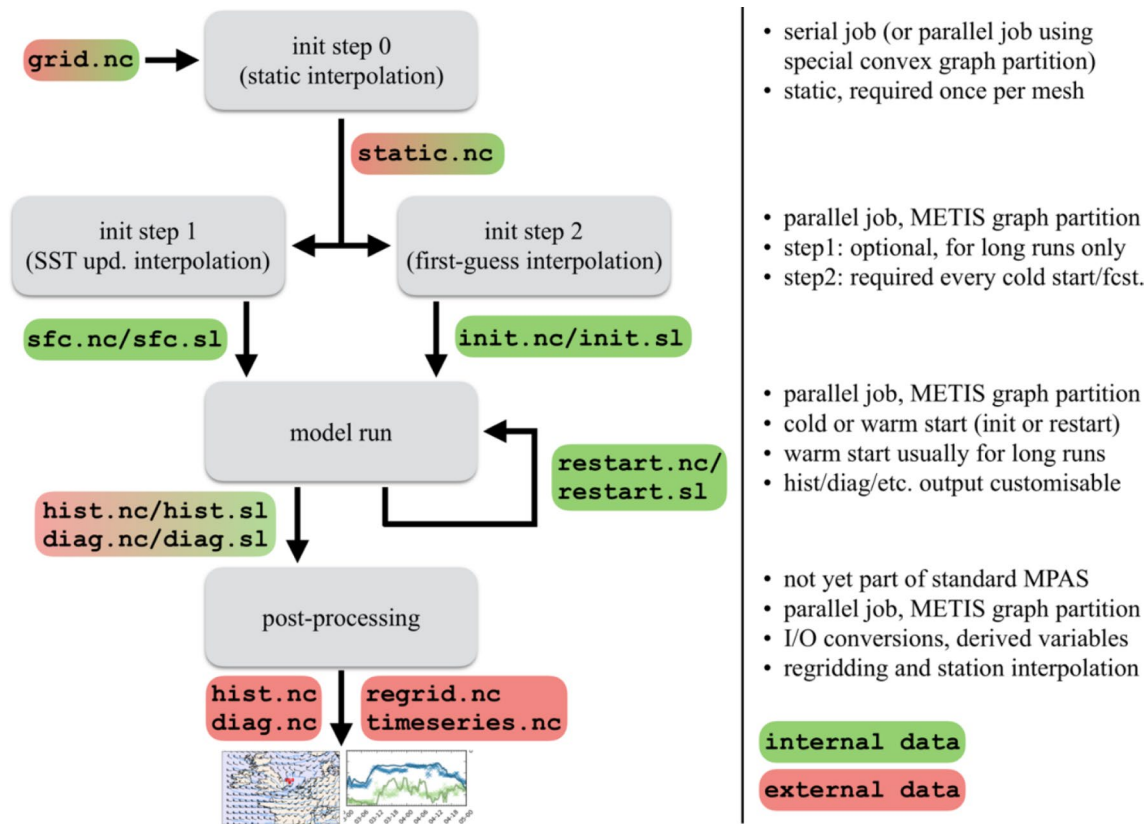


Fig. 13 Schematics of a typical MPAS model execution chain. The mesh file `grid.nc` contains the horizontal geometry of the mesh on a unit sphere. The static interpolation maps time-invariant data such as topography, vegetation cover, land-sea mask etc. onto the mesh and scales it to the actual radius of the Earth. The sea-surface temperature (SST) update interpolation creates regular boundary conditions to update SST and sea-ice from a forcing data set, usually only required for longer runs. The first guess interpolation maps a forcing data set (first-guess) onto the static file `static.nc`. While MPAS offers a restart

capability, this is usually not required in short-term weather forecasting. A recently developed post-processing core handles conversions between the file formats supported by MPAS (see text for details) as well as regridding to standard grids and station interpolation. The decomposition of the mesh for a given number of tasks is performed as a pre-processing step using the METIS graph partitioning software available at <http://glaros.dtc.umn.edu/gkhome/metis/metis/overview> (except for the static interpolation). For further details, see the MPAS Atmosphere Users' Guide available at <http://mpas-dev.github.io>

3 km mesh, and the 60–3 km mesh less than 1.3%. The generation of forecasts on large meshes also requires a certain amount of disk space for (temporarily) storing the model output. The amount of disk space required scales with the number of grid columns of the mesh, i.e. shows similar characteristics as the required computational resources discussed above. However, storage is more of an issue for longer runs such as seasonal predictions and climate projections than for short-term weather forecasting. For verification and model optimisation purposes in NWP, it should be sufficient to save the post-processed regridded or interpolated data or a subset of the global MPAS mesh (tools for working with such geographical subsets are currently under development at NCAR/LANL; M.G. Duda, priv. comm.), rather than keeping the full native MPAS output. The interested reader is referred to Heinzeller et al. (2016) and Heinzeller and Duda (2017) for further information on storage requirements and parallel I/O strategies.

The costs for generating a 36-h forecast with WRF are considerably lower. This is to be expected as the nested WRF domain contains only 374,544 grid columns, while the smallest of the global MPAS meshes contains 835,586 grid columns. However, removing this factor 2.2 from the equation, MPAS requires approximately five times as many computational resources as WRF. Several aspects should be noted here that contribute to the observed difference in computational requirements: (1) The operational WRF setup uses an adaptive time step that can range from 10 to 15 s, which commonly leads to a shorter time to delivery than the usual fixed time step in WRF (6 s per km grid spacing). In practise, the time step of WRF started at < 15 s, but quickly rose and stabilised at 15 s. The MPAS runs in this study use a conservative fixed time step of 15 s for all meshes, leading to generally the same time step for both models. (2) The operational 3 km WRF model is run in the MeteoGroup in-house cluster and has been optimised for best performance

Table 8 Computational costs of the different MPAS runs on SuperMUC

Mesh	Nodes	Tasks	Cells/task =: A	Initialisation in [s]	Time integration (1 h) in [s] =: B		B/A
					Dynamics + physics	File output	
3km*	1024	16,384	4000	694	1394	1081	0.62
3 km	1024	16,384	4000	128	1407	62	0.37
15–3 km	256	4096	1584	20	589	7	0.38
30–3 km	128	2048	632	14	232	7	0.38
30–3km**	128	2048	632	14	204	7	0.33
60–3 km	64	1024	816	10	296	4	0.37
60–3 km	128	2048	408	11	143	16	0.39
60–3 km	192	3072	272	12	96	8	0.38
60–3 km	256	4096	204	12	73	4	0.38
60–3 km	384	6144	136	16	56	5	0.47
60–3 km	512	8192	102	20	43	16	0.58

For the 3 km, 15–3 km and 30–3 km meshes, the configurations used to create the forecasts in this study are listed; for the 60–3 km mesh, 64 nodes were used. The 3km* setup uses PnetCDF (<https://trac.mcs.anl.gov/projects/parallel-netcdf>) for the file I/O, whereas all other setups use SIONlib (http://www.fzjuelich.de/ias/jsc/EN/Expertise/Support/Software/SIONlib/_node.html). The latter reduces the times required for model initialisation and file output and provides better scaling (see Heinzeller and Duda 2017). The 30–3 km** scaling results were obtained with a newer version of MPAS, based on release 5.1 with improvements of the hybrid parallelisation in the dynamical solver. The ratio of time integration (B) to cells per task (A; cells for which a core is “responsible”) in the last column is a measure for the parallel performance of the setup and shows that the realtime required for the time integration (column B) is similar for the various meshes employed in this study for similar numbers of cells per task (column A), as long as this value is larger than about 200 (using SIONlib for the file output). The total number of grid columns (cells) is 835,586 (60–3 km), 1,294,335 (30–3 km), 6,488,066 (15–3 km) and 65,536,002 (3 km)

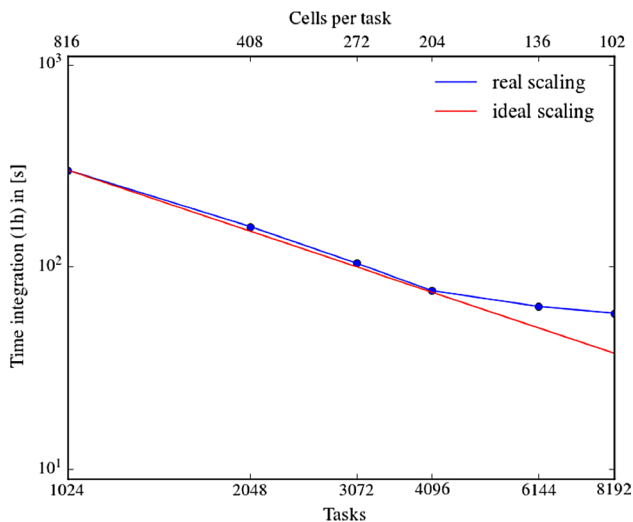


Fig. 14 Scaling of the time integration (see Table 8) for the 60–3 km grid on SuperMUC as a function of tasks (bottom) and cells per task (top; cells for which a core is “responsible”). A similar scaling behaviour is observed for the other meshes for the same numbers of cells per task (not displayed). For less than approximately 200 cells per task, the parallel performance of the time integration (dynamics, physics, file output) breaks down

on this system and for the given setup (e.g. the number of grid columns and vertical levels for this particular setup are fixed and specified during compile time, which leads to shorter runtimes; Wang et al. 2012). The MPAS performance was measured on.

SuperMUC using standard optimisations, i.e. without any further tuning of compiler flags. Using the latest version 5.1 of MPAS with performance improvements in the dynamical solver, in particular with respect to the hybrid MPI + OpenMP parallelisation, a decrease in runtime by 12% is found compared to the version used in this study (see Table 8). (3) The computational costs also depend on the chosen physics parameterisations (see Table 1). The WSM6 microphysics scheme employed by WRF is cheaper to run than the complex Thompson microphysics scheme employed by MPAS, which, for example, couples the cloud physics to the RRTMG radiation schemes and as such informs the radiation schemes about the microphysics—derived effective droplet sizes and optical depths (Thompson et al. 2016). It is however difficult to assess the difference this creates in run time, as the version of MPAS we used does not have support for WSM6, and the Thompson scheme is not coupled to other parameterisations in WRF. While MPAS uses the scale-aware Grell-Freitas scheme for convection, the cumulus parameterisation is switched off for WRF. In

Table 9 Required computational resources for an operational application of the different MPAS meshes employed in this study

36-h forecast scenario		Mesh	Nodes required	Cores required	Core-hours fcst
Time to solution [h]	4	3 km	3927	62,832	251,328
Ratio B/A	0.38	15–3 km	386	6176	24,704
Cells per task	1050	30–3 km	77	1232	4728
Cores per node (WRF)	8	60–3 km	50	800	3200
Cores per node (MPAS)	16	WRF 3 km	1070	70	280

The minimum number of nodes (tasks) on SuperMUC is constrained by the requested time to solution for a given forecast length (model integration time) and is calculated based on a B/A ratio of 0.38. The number of cells per task lies within the valid range for the assumed B/A ratio. For reference, the operational WRF setup at MeteoGroup is given, albeit the differences in time integration techniques (adaptive time step for WRF versus fixed time step for MPAS) and the different HPC systems and optimisations used (see text for details)

summary, this comparison between runtime between WRF and MPAS should be judged with care due to the differences in setup, especially regarding the number of vertical levels and parameterization settings. However, it can still be noted that the scaling of the MPAS model is impressive.

7 Discussion and conclusions

This study evaluates the MPAS global model at different resolutions and the WRF mesoscale model for three weather events within Europe, i.e. a gale, a föhn, and a convective event. The version of MPAS employed in this study is based on release 4.0 from late 2015 with several additions to the model, provided by NCAR and KIT. Numerous improvements were made in release 5.0 of MPAS from January 2017, in particular with respect to the scale-aware parameterisation of convection employed in this study (M.G. Duda, priv. comm.), but also in terms of the runtime performance of the model. WRF version 3.6.1 was used.

The WRF model performs well in all three cases, since it has many parameterisation options and this setup has been optimised for operational weather forecasting in Europe. For all three test cases, the tested MPAS meshes on average show a comparable performance to the operational WRF setup. Given that the WRF setup experienced years of optimisation while MPAS was used almost out of the box, this is an impressive result. However, one should note that in the interest of a fair comparison, WRF in this study was used without the data assimilation framework of MeteoGroup that is applied in operational forecasting. The use of a data assimilation with MPAS is still under development with some early fruitful results (Ha et al. 2017; Bullock et al. 2018), but this may pose limitations to regular users for employing MPAS in an operational setup straight away. Additionally, the WRF model uses 39 vertical layers, while the MPAS model uses 55. A test case using WRF with 55 vertical layers, similar to the MPAS setup, showed only minor differences compared to the 39-level WRF run (not shown).

The variable-resolution MPAS meshes show a similar forecast accuracy compared to the uniform 3 km mesh at much lower computational costs (1.3% for the 60–3 km mesh). For two of the three test cases (1: synoptic gale, 2: föhn effect), the more expensive MPAS meshes (15–3 km, 3 km) performed slightly better than the cheaper options (30–3 km, 60–3 km); for the remaining test case (3: organised convection), the cheaper options outperformed the more expensive meshes.

In general, the difference in forecasting accuracy between MPAS grid meshes is relatively small, compared to the differences caused by different initialisation times. Depending on the test case, an earlier or later initialisation leads to better results. The dependence of MPAS on the initial conditions is stronger than for WRF, where nudging techniques (not used here) and regular boundary updates influence the model run. In this study, a later initialisation closer to the event led to better results for test case 1. For test cases 2 and 3, we found a mixed bag with some aspects being more accurate in the early initialisation runs and some being more accurate in the later initialisation runs. In an operational setting, this aspect is of relatively small importance as the models are run at every initialisation time and often fed into a post processing system that employs state of the art statistical techniques.

The ability to produce high-quality forecasts with MPAS on variable-resolution meshes is encouraging as it enables an operational application on current computational facilities. While a uniform MPAS 3 km would enable weather forecast providers to produce high quality forecasts everywhere on the globe, such a setup is currently not feasible, since it requires tremendous computational resources. However, with exascale being projected for the end of the decade, the supreme parallel scaling of MPAS and further optimisations of the model (Heinzeller et al. 2017), global convection-resolving forecasts will be possible in the next years, albeit this will be limited to a small number of operational centres with the necessary computational resources.

The computational requirements to generate a forecast with MPAS are higher than for WRF, which is partly due

to the larger number of grid columns for a global MPAS mesh, compared to a nested WRF grid. For similar numbers of grid columns, the version of MPAS employed here requires approximately five times the computational resources of WRF to generate a forecast within the same time, although this comparison is hampered by the different number of vertical levels and the different physics parameterisations. Further, this number will likely be reduced in an operational setting, where similar efforts are made to optimise the model performance on the available system and for the particular mesh, and with newer versions of the model: using the latest release of MPAS version 5.1 leads to 12% shorter runtimes and shows the potential for further speed up (Heinzeller et al. 2017). Also, in tests of WRF and MPAS using the same setup (physics schemes, vertical levels) and optimisation with the latest development version of MPAS, the runtime performance was only 40% slower than that of WRF (M.G. Duda, priv. comm.).

While the variable-resolution MPAS meshes employed in this study contained only one circular area of refinement, it is possible to design multiple areas of refinement in—to some extent—arbitrary shape and size. This capability is highly attractive for global weather service providers as it allows them to run a single global MPAS forecast rather than multiple nested WRF forecasts scattered across the globe. In summary, an application of MPAS in real-time forecasting is feasible and a promising option for weather service providers such as MeteoGroup, although more research is needed in particular with respect to data assimilation and post-processing tools that are able to digest large amounts of output on unstructured grids.

Furthermore, the WRF model is driven by a global analysis. Therefore, the MPAS system could effectively be the source of initial and boundary conditions, as it is a global model. However, the quality of the EMWF fields are outstanding and are therefore preferred as driving fields in this study. Comparing the forecast quality of a regional model, where the source of the forecast fields differs between MPAS and for example ECMWF, are topics of another study. In addition, it would ignore the power of MPAS allowing grid refinement, of which we want to evaluate the impact here.

In some global models with irregular grid meshes, grid imprinting can occur due to the presence of pentagons in the otherwise hexagonal grid. However, this is only an issue when regular hexagons are used to fill the globe. In this case, only hexagons were used. We did not study the occurrence or effect of grid imprinting for the current case studies. It is worth noting that Park et al. (2013) did not find any impact of grid imprinting in their case.

Acknowledgements The authors thank KNMI and ECMWF for making their data available. We thank Michael Duda for fruitful discussions

on several aspects addressed in this paper. GJ. Steeneveld acknowledges funding from NWO grant 864.14.007.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alam MM (2014) Impact of cloud microphysics and cumulus parameterization on simulation of heavy rainfall event during 7–9 October 2007 over Bangladesh. *J of Earth Syst Sci* 123(2):259–279. <https://doi.org/10.1007/s12040-013-0401-0>
- Beljaars ACM (1987) The Influence of sampling and filtering on measured wind gusts. *J Atmos Oceanic Technol* 4:613–626
- Bukovsky MS, Karoly DJ (2009) Precipitation simulations using WRF as a nested regional climate model. *J Appl Meteorol Clim* 48:2152–2159. <https://doi.org/10.1175/2009JAMC2186.1>
- Bullock Jr. OR, Foroutan, H, Gilliam RC, Herwehe JA (2018) Adding four-dimensional data assimilation by analysis nudging to the model for prediction across scales—atmosphere (version 4.0), *Geosci Model Dev*, 11: 2897–2922
- Case JL, Kumar SV, Srikishen J, Jedlovec GJ (2011) Improving numerical weather predictions of summertime precipitation over the Southeastern United States through a high-resolution initialisation of the surface state. *Wea Forecast* 26(6):785–807. <https://doi.org/10.1175/2011WAF2222455.1>
- Crétat J, Pohl B, Richard Y, Drobinski P (2012) Uncertainties in simulating regional climate of Southern Africa: sensitivity to physical parameterisations using WRF. *Clim Dyn* 38(3–4):613–634. <https://doi.org/10.1007/s00382-011-1055-8>
- Cuo L, Pagano TC, Wang QJ (2011) A review of quantitative precipitation forecasts and their use in short- to medium-range streamflow forecasting. *J Hydrometeor* 12(5):713–728. <https://doi.org/10.1175/2011JHM1347.1>
- Davis CA, Ahijevych DA, Wang W, Skamarock WC (2016) Evaluating medium-range tropical cyclone forecasts in uniform- and variable-resolution global models. *Mon Wea Rev* 144:4141–4160. <https://doi.org/10.1175/MWR-D-16-0021.1>
- Ducrocq V, Ricard D, Lafore JP, Orain F (2002) Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Weather Forecast* 17(6):1236–1256
- Ebert EE, Damrath U, Wergen W, Baldwin ME (2003) The WGENE assessment of short-term quantitative precipitation forecasts. *Bull Amer Meteor Soc* 84(4):481–492 + 429. <https://doi.org/10.1175/BAMS-84-4-481>
- Ebert EE, Janowiak JE, Kidd C (2007) Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull Amer Meteor Soc* 88(1):47–64. <https://doi.org/10.1175/BAMS-88-1-47>
- Eden JM, Widmann M, Grawe D, Rast S (2012) Skill, correction, and downscaling of GCM-simulated precipitation. *J Clim* 25(11):390–3984. <https://doi.org/10.1175/JCLI-D-11-00254.1>
- Ek MB, Mitchell KE, Lin Y, Rogers E, Grunmann P, Koren V, Gayno G, Tarpley JD (2003) Implementation of Noah land surface model advances in the national centers for environmental prediction operational mesoscale eta model. *J Geophys Res-Atmos* 108(22):GCP12–G1

- Flaounas E, Bastin S, Janicot S (2011) Regional climate modelling of the 2006 West African monsoon: sensitivity to convection and planetary boundary layer parameterisation using WRF. *Clim Dyn* 36(5–6):1083–1105. <https://doi.org/10.1007/s00382-010-0785-3>
- Fowler LD, Skamarock WC, Grell GA, Freitas SR, Duda MG (2016) Analyzing the Grell–Freitas convection scheme from hydrostatic to nonhydrostatic scales within a global model. *Mon Weather Rev* 144:2285–2306. <https://doi.org/10.1175/MWR-D-15-0311.1>
- Gadian A, Blyth A, Bruyere CL, Burton R, Done J, Groves J, Holland G, Mobbs S, Pozo JT, Tye M, Warner J (2018) A case study of possible future summer convective precipitation over the UK and Europe from a regional climate projection. *Int J Climatol* 38:2314–2324. <https://doi.org/10.1002/joc.5336>
- Gomes JL, Chou SC (2010) Dependence of partitioning of model implicit and explicit precipitation on horizontal resolution. *Meteorol Atmos Phys* 106(1):1–18
- Grell GA, Freitas SR (2013) A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos Chem Phys* 14(10):5233–5250. <https://doi.org/10.5194/acp-14-5233-2014>
- Ha S, Snyder C, Skamarock WC, Anderson J, Collins N (2017) Ensemble kalman filter data assimilation for the model for prediction across scales (MPAS). *Mon Weather Rev* 145:4673–4692
- Heinzeller D, Duda MG (2017) Preparing for exascale: A global 2 km atmospheric scaling experiment with the model for prediction across scales (MPAS). In: Technical report JUQUEEN extreme scaling workshop 2017. Brömmel D, Frings W, Wylie BJN (Eds), held January 23–25 2017 in Jülich, Germany. <http://hdl.handle.net/2128/13977>. Accessed 29 Sept 2016
- Heinzeller D, Duda MG, Kunstmann H (2017) Preparing for exascale: towards convection-permitting, global atmospheric simulations with the model for prediction across scales (MPAS), in prep. for submission to Geoscientific Model Development
- Heinzeller D, Duda M, Kunstmann H (2016) Towards convection-resolving, global atmospheric simulations with the model for prediction across scales (MPAS) v3.1: an extreme scaling experiment. *Geosci Model Dev* 9:77–110
- Hong SY (2004) Comparison of heavy rainfall mechanisms in Korea and the central US. *J Meteorol Soc Jpn* 82:1469–1479
- Hong SY, Lim JOJ (2006) The WRF single-moment 6-class microphysics scheme (WSM6). *J Korean Meteor Soc* 42(2):129–151
- Iacono MJ, Delamere JS, Mlawer EJ, Shephard MW, Clough SA, Collins WD (2008) Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *J Geophys Res* 113:D13103. <https://doi.org/10.1029/2008JD009944>
- Jankov I, Gallus WA Jr, Shaw B, Koc SE (2005) The impact of different WRF model physical parameterisations and their interactions on warm season MCS rainfall. *Weather Forecast* 20(6):1048–1060
- KNMI (2017) KNMI Data center NOAA-AHVR image products. http://projects.knmi.nl/datacenter/satellite_earth_observations/NOAA/archive/. Accessed 14 March 2016
- KNMI (2015) Onweer 30–31 augustus. <https://www.knmi.nl/kennis-en-datacenter/achtergrond/onweer-30-31-augustus-2015>. Accessed 12 Sept 2016
- Klein C, Heinzeller D, Bliedernicht J, Kunstmann H (2015) Variability of West African monsoon patterns generated by a WRF multi-physics ensemble. *Clim Dyn* 45(9–10):2733–2755. <https://doi.org/10.1007/s00382-015-2505-5>
- Krikken F, Steeneveld GJ (2012) Modeling the Reintensification of tropical storm erin (2007) over Oklahoma: understanding the key role of downdraft formulation, *Tellus A64*, 17417, <https://doi.org/10.3402/tellusa.v64i0.17417>
- Kunstmann H, Stadler C (2005) High resolution distributed atmospheric-hydrological modelling for Alpine Catchments. *J Hydrol* 314:105–124
- Landu K, Leung R, Hagos S, Vиноj V (2014) The dependence of ITCZ structure on model resolution and dynamical core in aquaplanet simulations. *J Clim* 27:2375–2385. <https://doi.org/10.1175/JCLI-D-13-00269.1>
- Laprise R, Kiornic D, Rapačić M, Šeparović L, Leduc M, Nikiema O, Luca AD, Diaconescu E, Alexandru A, Lucas-Picher P, de Elía R, Caya D, Biner S (2012) Considerations of domain size and large-scale driving for nested regional climate models: Impact on internal variability and ability at developing small-scale details. In: Berger A, Mesinger F, Sijacki D (eds) *Climate change: inferences from paleoclimate and regional aspects*. Springer, Vienna, pp 181–199. https://doi.org/10.1007/978-3-7091-0973-1_14
- Leduc M, Laprise R (2009) Regional climate model sensitivity to domain size. *Clim Dyn* 37(6):833–854
- Leduc M, Laprise R, Moretti-Poisson M, Morin JP (2011) Sensitivity to domain size of mid-latitude summer simulations with a regional climate model. *Clim Dyn* 37:343–356. <https://doi.org/10.1007/s00382-011-1008-2>
- Lim KSS, Hong SY, Yoon JH, Han J (2014) Simulation of the summer monsoon rainfall over East Asia using the NCEP GFS cumulus parameterisation at different horizontal resolutions. *Wea Forecast* 29(5):1143–1154. <https://doi.org/10.1175/WAF-D-13-00143.1>
- Nakanishi (2010) Improvement of the Mellor–Yamada turbulence closure model based on largeEddy simulation data. *Bound-Layer Meteorol* 99:349–378
- Park S, Skamarock WC, Klemp JB, Fowler LD, Duda MG (2013) Evaluation of global atmospheric solvers using extensions of the Jablonowski and Williamson baroclinic wave test case. *Mon Wea Rev* 141:3116–3129
- Pei L, Moore N, Zhong S, Luo L, Hyndman DW, Heilman WE, Gao Z (2014) WRF model sensitivity to land surface model and cumulus parameterisation under short-term climate extremes over the southern Great Plains of the United States. *J Clim* 27(20):7703–7724. <https://doi.org/10.1175/JCLI-D-14-00015.1>
- Pennely C, Reuter G, Flesh T (2014) Verification of the WRF model for simulating heavy precipitation in Alberta. *Atmos Res* 135:172–192. <https://doi.org/10.1016/j.atmosres.2013.09.004>
- Pilon R, Zhang C, Dudhia J (2016) Roles of deep and shallow convection and microphysics in the MJO simulated by the model for prediction across scales. *J Geophys Res Atmos* 121:10575–10600. <https://doi.org/10.1002/2015JD024697>
- Price E, Mielikainen J, Huang M, Huang B, Huang HLA, Lee T (2014) GPU-accelerated longwave radiation scheme of the rapid radiative transfer model for general circulation models (RRTMG). *IEEE J Sel Top Appl Earth Obs Remote Sens* 7(8):3660–3667
- Schär C, Leuenberger D, Fuhrer O, Lüthi D, Girard C (2002) A new terrain-following vertical coordinate formulation for atmospheric prediction models. *Mon Wea Rev* 130:2459–2480
- Schwartz Cs, Kain JS, Weiss SJ, Xue M, Bright DR, Kong F, Thomas KW, Levit JJ, Coniglio MC, Wandishin MS (2010) Toward improved convection-allowing ensembles: model physics sensitivities and optimising probabilistic guidance with small ensemble membership. *Wea Forecast* 25(1):263–280. <https://doi.org/10.1175/2009WAF2222267.1>
- Skamarock WC, Klemp JB (2008) A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J Comp Phys* 227(7):3465–3485. <https://doi.org/10.1016/j.jcp.2007.01.037>
- Skamarock WC, Klemp JB, Duda MG, Fowler LD, Park SH, Ringler TD (2012) A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tessellations and C-grid staggering. *Mon Weather Rev* 240:3090–3105
- Smiatek G, Kunstmann H, Werhahn J (2012) Implementation and performance analysis of a high resolution coupled numerical weather and river runoff prediction model system for an Alpine catchment.

- Environ Model Softw 38:231–243. <https://doi.org/10.1016/j.envsoft.2012.06.001>
- Staneva J, Alari V, Breivik Ø, Bidlot JR, Mogensen K (2017) Effects of wave-induced forcing on a circulation model of the North Sea. *Ocean Dyn* 67(1):81–101. <https://doi.org/10.1007/s10236-016-1009-0>
- Steenefeld GJ, Ronda RJ, Holtslag AAM (2015) The challenge of forecasting the onset and development of radiation fog using mesoscale atmospheric models. *Bound-Layer Meteorol* 154:265–289
- Steppeler J, Bitzer H, Minotte M, Bonaventura L (2002) Nonhydrostatic atmospheric modeling using a z-coordinate representation. *Mon Weather Rev* 130:2143–2149
- Thompson G, Rasmussen RM, Manning K (2004) Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: description and sensitivity analysis. *Mon Weather Rev* 132(2):519–542
- Wadey MP, Brown JM, Haigh ID, Dolphin, Wisse P (2015) Assessment and comparison of extreme sea levels and waves during the 2013/14 storm season in two UK coastal regions. *Nat Hazards Earth Syst Sci* 15(10):2209–2225. <https://doi.org/10.5194/nhess-15-2209-2015>
- Wang W, Beezley C, Duda M (2012) WRF ARW V3: user's guide. National Center for Atmospheric Research, Boulder
- Wooten A, Bowden JH, Boyles R, Terando A (2016) The sensitivity of WRF downscaled precipitation in puerto rico to cumulus parameterization and interior grid nudging. *J Appl Meteor Climatol* 55:2263–2281

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.