

Sensitivity analysis of mesoscale simulations to physics parameterizations : a case study of storm Ciara over the Belgian North Sea using WRF-ARW

Adithya Vemuri^{1,2,3}, Sophia Buckingham¹, Wim Munters¹, Jan Helsen², and Jeroen van Beeck¹

¹Department of Environmental and Applied Fluid Dynamics, von Karman Institute for Fluid Dynamics, Waterloosesteenweg 72, 1640 Sint-Genesius-Rode, Belgium

²Department of Mechanical Engineering, Vrije Universiteit Brussel, Boulevard de la Plaine 2, 1050 Ixelles, Belgium

³SIM vzw, Technologiepark 48, 9052 Zwijnaarde, Belgium

Correspondence: Adithya Vemuri (adithya.vemuri@vki.ac.be)

Abstract.

The Weather, Research and Forecasting (WRF) model ~~includes~~ offers a multitude of physics parameterizations to ~~account for atmospheric dynamics and interactions such as turbulent fluxes within the planetary boundary layer (PBL), long and short wave radiation, hydrometeor representation in microphysics, cloud ensemble representation in cumulus, amongst others.~~ A study and
5 analyze the different atmospheric processes and dynamics that are observed in the Earth's atmosphere. However, the suitability of a WRF model setup is known to be highly sensitive to the type of weather phenomena and the type and combination of physics parameterizations. A multi-event sensitivity analysis is conducted ~~in order to identify the optimal WRF-physics set-up and impact of temporal resolution of re-analysis dataset for the event of sudden changes in wind direction that can become challenging for reliable wind energy operations. In this context, Storm Ciara has been selected as a case study to investigate the~~
10 ~~influence of a broad combination of different interacting physics schemes on quantities of interest that are relevant for energy yield assessment. Of particular relevance to fast transient weather events, two different temporal resolutions (1-hourly and 3-hourly) of the lateral boundary condition's re-analysis dataset, ERA5, are considered~~ to identify general trends and suitable WRF physics setups for 3 extreme weather events identified to be potentially harmful for the operation and maintenance of
15 a low-pressure system on 24 December 2020, and a trough passage on 27 June 2020. 12 WRF simulations per event are performed to study the effect of the update interval of lateral boundary conditions and different combinations of physics parameterizations (PBL, cumulus, and microphysics). Specifically, the update interval of ERA5 lateral boundary conditions is varied between hourly and 3-hourly. Physics parameterizations ~~considered in this study include: two are varied between 3 PBL schemes (MYNN2.5-MYNN, scale-aware Shin-Hong, and scale-aware Shin-Hong-PBL), four Zhang), 4 cumulus schemes~~
20 ~~(Kain-Fritsch, Grell-Devenyi, and Grell-Dévényi, scale-aware Grell-Freitas, and multi-scale Kain-Fritsch,) and three), and 3 microphysics schemes (WSM5, Thompson, and Morrison) coupled with two geospatial configurations for WRF simulation domains. The resulting WRF predictions are assessed by comparison to observational RADAR reflectivity data on precipitation. In addition, SCADA data on~~ The simulated wind direction and wind speed are compared qualitatively and quantitatively (using

MAE) to operational SCADA data. Overall, a definitive best-case setup common to all three events is not identified in this study. For wind direction and wind speed from an offshore wind farm located in the Belgian North Sea is considered to assess modeling capabilities for local wind behavior at farm level. For precipitation, results are shown to be very sensitive to model setup, but no clear trends can be observed. For wind-related variables on the other hand, results show a definite improvement in accuracy when both, the best-case setups are identified to employ scale-aware cumulus and PBL parameterizations are used in combination with 1-hourly temporal resolution reanalysis data and extended domain sizes. PBL schemes. These are most often driven by hourly update intervals of lateral boundary conditions as opposed to 3-hourly, although it is only in the case of storm Ciara that significant differences are observed. Scale-aware cumulus schemes are identified to produce better results when combined with scale-aware PBL schemes, specifically for Storm Ciara and the trough passage cases. However, for the low-pressure system case this trend is not observed. No clear trend in utilizing higher-order microphysics parameterization considering the combinations of WRF setups in this study is found in all cases. Overall, the combination of PBL, cumulus, and microphysics schemes is found to be highly sensitive to the type of extreme weather event. Qualitatively, precipitation fields are found to be highly sensitive to model setup and the type of weather phenomena.

1 Introduction

Extreme weather phenomena such as low-level jets, sudden-fast changes in wind direction, extreme wind shear (Kalverla et al., 2017; Aird et al., 2021), wind ramps (Gallego-Castillo et al., 2015), and storms (Solari, 2020) are capable of causing severe dynamic loading on wind turbine components-turbines (Negro et al., 2014; AbuGazia et al., 2020; Chi et al., 2020). Furthermore, precipitation associated to-with these phenomena can lead to early blade degradation through leading-edge erosion (Law and Koutsos, 2020). As such, these extreme-weather-events-Extreme Weather Events (EWE) play a significant role in the wind-turbine's-wind turbine operational lifetime and must be considered at the design stage to ensure that-ultimate-loads-are-not-exceeded-and-fatigue-requirements-are-met. Furthermore, such events may safe estimates of ultimate and fatigue loading. Such events may also cause sudden changes in power production leading to grid imbalance and economic losses. Accurate Therefore, accurate modeling and forecasting of such EWE is hence crucial to tackle these challenges in view of current and future expansion of both onshore and offshore wind energy. Numerical weather prediction farms. Typically, Numerical Weather Prediction (NWP) models provide a promising approach to help identify problematic weather events and to predict their occurrence through operational forecasting (Bauer et al., 2015). Among different available NWP codes, the open-source Weather, are utilized to identify, study, and analyze such extreme weather phenomena. Recent developments in NWP models pave the way towards high resolution weather forecasts, thus enabling operational use for wind energy applications (Dudhia, 2014; Bauer et al., 2015). This study utilizes the public domain Weather Research and Forecasting - Advanced Research WRF (WRF-ARW) model developed by NCAR/NOAA (?) is commonly used for its ability to represent the various interacting the National Center for Atmospheric Research (Skamarock et al., 2019; Powers et al., 2017). The WRF model represents a multitude of atmospheric processes and dynamics of the atmosphere-such as the distribution of fluxes within the planetary-boundary-layer-Planetary Boundary Layer (PBL), the determination of cloud ensembles and

compensating subsidence for convective /cumulus systems, ~~the accurate representation of evolving hydrometeor species, solar irradiation, land-surface interactions and heat and moisture fluxes in the surface layer (SL). An expanse and the evolution of hydrometeor species. Therein, an array of physics parameterizations and options are available in WRF to represent the influence of these phenomena on local weather systems. However, predictions model parameters are available to adequately represent a local weather system. Nonetheless, WRF simulations~~ are found to be highly sensitive ~~on the selection of these sub-grid scale models to the type and combination of physics schemes,~~ the location and the type of weather event, ~~the lateral boundary conditions used to drive the flow, and the simulation domain configuration~~ and the Lateral Boundary Conditions (LBC).

Sensitivity analyses are typically conducted to identify the optimal combination of ~~physics schemes in the event of a specific type of weather system over a given area~~ physics schemes for a specific location (see, e.g., Efstathiou et al. 2013; Santos-Alamillos et al. 2013; Kala et al. 2015). ~~To date, this~~ This type of investigation has not been performed ~~over for~~ the Belgian North Sea, ~~nor has any previous study.~~ Furthermore, to the authors' best knowledge ~~looked at how the extreme weather event is experienced from a wind farm perspective through comparison with operational SCADA data. This work aims to tackle this challenge by considering,~~ no previous studies have looked at potentially harmful EWE from a wind farm perspective as experienced by the machines themselves. Therefore, this sensitivity analysis aims to address this gap in research. The analysis presented in this paper assesses the impact of a wide range of ~~interacting physical parameterizations,~~ more specifically on cumulus, microphysics and PBL schemes. ~~Furthermore, we assess the need for high temporal resolution mesoscale forcing data and extended numerical domains for the prediction of wind~~ physics parameterizations for PBL, cumulus and precipitation quantities of interest for wind farm design and operation. microphysics, and length of the update interval of LBC on the simulated wind direction and wind speed.

~~The cumulus, microphysics and PBL parameterizations defined in WRF follow~~ WRF physics parameterizations for PBL, cumulus, and microphysics comprise a multitude of large-scale and sub-grid scale modeling techniques. ~~For PBL and cumulus, these are primarily divided into scale-aware and non-scale-aware non-scale-aware parameterizations. The scale-aware schemes promise to better reproduce parameterizations aim to better represent convective and turbulent mixing effects in fluxes at the~~ so-called gray-zone gray-zone resolutions, i.e. for high-resolution simulation grids, for refined horizontal grid spacings which are on the verge of allowing explicitly partial resolution of these effects fluxes rather than fully parameterizing them (Wyngaard, 2004; Hong and Dudhia, 2012). ~~The cumulus parameterizations represent the ensemble effects of convective clouds following paragraphs briefly discuss the state-of-the-art physics parameterizations of PBL, cumulus, and microphysics.~~

Concerning the parameterization of boundary-layer turbulence, traditional PBL schemes rely on the assumption of horizontal homogeneity to redistribute surface fluxes vertically within the atmospheric boundary layer. However, for horizontal grid spacings of around 1 km or finer, three-dimensional atmospheric turbulence becomes partially resolved, violating this basic assumption employed by classical 1D PBL schemes. The gray-zone modeling challenge for PBL turbulence has led to the development of scale-aware PBL schemes which partially resolve turbulent mixing at gray-zone resolutions as a function of larger-scale processes and conditions by formulating the the grid spacing. This work considers 3 PBL parameterizations: the non-scale-aware 1D Mellor-Yamada-Nakanishi-Niino (MYNN) scheme, the scale-aware 1D Shin-Hong (SH) scheme, and the scale-aware 3D Zhang scheme. The MYNN PBL scheme (Nakanishi and Niino, 2006) is a 1D turbulence kinetic energy

prediction scheme that solves for a vertical eddy viscosity profile in a grid column. The SH PBL scheme (Shin and Hong, 2015) is a scale-aware 1D diagnostic non-local scheme representing non-local transport by large eddies in the atmospheric boundary layer. The SH PBL scheme modifies the YSU PBL scheme (Hong et al., 2006) for sub-kilometer transition scales by reducing the strength of the non-local term with decreasing horizontal grid spacing, assuming gradual resolution of the largest eddies. The Zhang 3D PBL scheme (Zhang et al., 2018) extends the 3D turbulent kinetic energy based closure by Deardorff (Deardorff, 1980) to gray-zone resolutions, using partitioning functions derived from a reference large eddy simulation. While the SH PBL scheme has been found to outperform conventional PBL formulations for desert convective boundary layers (Xu et al., 2018) and for the western Great Plains of the United States (Doubrawa and Muñoz-Esparza, 2020), its interaction with cumulus and microphysics options is yet to be tested for extreme weather in coastal environments featuring strong interaction between PBL and convective cumulus processes.

The cumulus parameterizations represent the ensemble effects of convective clouds with statistical effects of moist convection and convective rainfall within a grid column. These grid column. Cumulus schemes are further divided into mass-flux type and adjustment type schemes. The mass-flux type schemes convert aim to minimize the convective available potential energy from a single grid column defined in WRF within a grid column by translating it into compensating subsidence. For example, a combination of vertical advection, moisture, and temperature. The current work considers the Kain-Fritsch (KF), the multi-scale Kain-Fritsch (msKF), the Grell-Dévényi 3D ensemble (GD-3D), and the scale-aware Grell-Freitas (GF) cumulus schemes, to evaluate the performance of WRF across the convective gray-zone. The KF cumulus scheme (Kain, 2004) is a commonly used one-dimensional commonly used 1D mass-flux cumulus-type scheme that considers deep and shallow convection and. The scheme includes hydrometeor detrainments from clouds, rain, ice, and snow. The scheme is designed to run at horizontal resolutions coarser than a horizontal grid spacing of 25 km. The multi-scale Kain-Fritsch (msKF) cumulus parameterization and coarser. The msKF cumulus scheme (Zheng et al., 2016) updates the original KF parameterization by introducing scale-aware parameterized cloud dynamics based on a dynamic length scale in order to improve the prediction accuracy at higher horizontal resolutions below KF cumulus scheme to convective gray-zone resolutions at horizontal grid spacings of 10 km. The Grell-Dévényi 3D ensemble (and coarser. The GD-3D) cumulus parameterization cumulus scheme (Grell and Dévényi, 2002) relies on combining ensemble and data assimilation techniques to represent the local convection and provides more tunable adjustable parameters for further calibration of the model. The Grell-Freitas (GF) cumulus parameterization scheme. The GF cumulus scheme (Grell and Freitas, 2014) is an adjustment type parameterization that explores; redistribution of redistributes compensating subsidence derived from GD-3D to neighbouring neighboring grid cells using distribution functions a Gaussian distribution function and adapts the scale-aware parameterization cloud representations from Arakawa et al. (2011). The GF cumulus parameterization was scheme is designed and tested for a horizontal resolution horizontal grid spacings of 5 km. The and coarser. A study by Jeworrek et al. (2019) highlights how crucial the choice in cumulus parameterization is for obtaining accurate WRF predictions of precipitation patterns, in particular when going from parameterized to resolved convective scales. In this work, the KF, msKF, GD-3D and GF schemes are considered in order to evaluate the performance of WRF across the the importance of choosing an appropriate cumulus parameterization to accurately represent precipitation, particularly in the convective gray-zone transition.

~~The microphysical parameterization is an emulation of processes for~~ Microphysics parameterizations emulate the processes of moisture removal from the atmosphere ~~via by~~ modeling hydrometeor distributions based on thermodynamic and kinematic fields defined in ~~the model. The~~ WRF. These schemes determine the spatial distribution of precipitation and vertical distribution of hydrometeor mass and latent heat. ~~Most commonly used microphysical~~ The most commonly used microphysics schemes are the so-called bulk schemes, ~~in which~~. These constitute a mathematical distribution ~~for of~~ hydrometeor number concentration versus particle size ~~is defined using using either a~~ negative exponential or ~~gamma distributions~~. Microphysical a gamma distribution. Bulk type microphysics parameterizations are further divided by ~~their complexity and tunable parameters such as~~ considered moments of distributions (single, double, etc), and intercepts or slope parameters for the order of complexity and number of tunable parameters, which define the moments and intercepts used by the aforementioned distributions. The microphysics schemes considered ~~in for~~ this sensitivity analysis are the ~~WSM5 single-moment~~ WRF Single-Moment 5-class (i.e. with scheme (WSM5) (Hong et al., 2004) representing 5 hydrometeor species) microphysics parameterization (Hong et al., 2004) classes of hydrometeor species, the Thompson single-moment (except ice) 6-class scheme (Thompson et al., 2008), and the Morrison double-moment 6-class scheme (Morrison et al., 2009). In this respect, Hong and Lim (2006) illustrated the advantages in including a greater number of hydrometeor species ~~for microphysical representations and a better prediction of~~ precipitation levels in microphysical representations to better predict precipitation fields. Similar results were observed ~~in a~~ recent study by Jeworrek et al. (2019), calling for microphysics parameterizations with greater fidelity in hydrometeor representation. The study finds that higher-order microphysics schemes, such as ~~the~~ Morrison and Thompson ~~schemes, and~~, in combination with scale-aware cumulus parameterizations, such as multi-scale KF and GF schemes, ~~in order to more~~ accurately reproduce precipitation.

~~Concerning the parameterization of boundary layer turbulence, mesoscale NWP models employ PBL schemes that rely on the horizontal homogeneity assumption to redistribute surface fluxes vertically within the atmospheric boundary layer. However, for grid spacings that are fine enough (≈ 1 km), three-dimensional atmospheric turbulence becomes partially resolved, which violates the basic assumption employed by classical one-dimensional PBL schemes. The gray-zone modeling challenge~~ for PBL turbulence has led to the development of scale-aware PBL schemes which, as opposed to non-scale-aware formulations, partially resolving turbulent mixing at gray-zone resolutions depending on grid size. This work considers two PBL parameterizations: a non-scale-aware Mellor-Yamada-Nakanishi-Niino (MYNN) scheme (Nakanishi and Niino, 2006) and a scale-aware Shin-Hong scheme (Shin and Hong, 2015). The MYNN PBL scheme is a one-dimensional turbulence kinetic energy prediction scheme that solves for a vertical eddy viscosity profile in a grid column considering, among others, buoyancy and shear production, boundary layer stability and vertical mixing. On the other hand, Shin-Hong is a scale-aware one-dimensional diagnostic non-local PBL scheme representing non-local transport by large eddies in the boundary layer. The Shin-Hong scheme modifies the YSU PBL scheme (Hong et al., 2006) for sub-kilometer transition scales (1 km down to 200 m) by reducing the strength of the non-local term with decreasing grid size, assuming gradual resolution of the largest eddies. While it has been found to outperform conventional PBL formulations for desert convective boundary layers (Xu et al., 2018) and for the XPIA study in the western Great Plains of the United States (Doubrawa and Muñoz-Esparza, 2020), its interaction with

~~cumulus and microphysics options is yet to be tested for extreme weather in coastal environments featuring strong interaction between PBL and microphysical processes.~~

In the context of offshore wind energy applications, various sensitivity studies have been conducted with the aim of determining a universal ~~"best" case WRF setup to assess local wind resources~~ best-case WRF setup for assessing the local weather systems (Hahmann et al., 2015; Giannakopoulou and Nhili, 2014; Carvalho et al., 2012). ~~The~~ Therein, the literature presents equivocal ~~results findings~~ from a multitude of sensitivity analyses conducted at various locations around the planet, ~~indicating a high dependency on the physics combinations and lateral boundary conditions used. Concerning the estimation of wind energy production compared to measured wind data, Hahmann et al. (2015) illustrating a strong dependence of WRF simulations to the type and combination of physics parameterizations, the initial and LBC, the horizontal and vertical grid spacing, and the location and type of weather phenomenon. For instance, comparing wind power production to observational data, Hahmann et al. (2015) study the long-term sensitivity of simulated WRF offshore climatology evaluated against wind LiDAR observations, indicating a strong sensitivity to PBL parameterizations and the spin-up period, and an insensitivity to global reanalysis and vertical resolution of the model. Carvalho et al. (2014), for offshore and onshore areas in the Iberian Peninsula, indicated grid spacing considered in the WRF model. Similarly, Carvalho et al. (2014) indicate a close dependency of PBL and SL parameterizations with on PBL and surface layer parameterizations studying different physics combinations, favoring better reproduction of different prognostic variables that may lead to increased accuracy depending on the prognostic variables of interest.~~ Cunden et al. (2018) performed a sensitivity analysis considering different combinations of non-scale-aware ~~cumulus, PBL and microphysics schemes~~ PBL, cumulus, and microphysics parameterizations (despite kilometer-range ~~resolutions grid spacing~~) for the ~~Island island~~ of Mauritius under clear and extreme weather (~~eyelonic and anti-eyelonic~~), ~~and were~~. The study was able to identify a ~~best case WRF setup~~ best-case WRF setup suitable for accurately simulating both cases. In contrast, ~~a similar~~ the study by Islam et al. (2015) for the Haiyan tropical cyclone over the west Pacific Ocean ~~concluded in no particular~~ did not identify a suitable combination of WRF physics to best reproduce the extreme weather event. ~~For~~ Similarly, for the European continent, studies by ~~García-Díez et al. (2013); Stergiou et al. (2017); Mooney et al. (2013) have conducted year-long and/or~~ García-Díez et al. (2013), Stergiou et al. (2017), and Mooney et al. (2013) have conducted long-term sensitivity ~~analyzes analyses~~ indicating a wide ~~spread of possible array of possible combinations of~~ physics parameterizations depending on the type of ~~local weather event, season and time-lapse considered~~ weather phenomenon, the season, and the time period to simulate within the diurnal cycle.

~~As discussed above, the~~ The optimal selection of WRF physics parameterizations remains an important ~~open challenge for accurate wind and weather modeling~~ and open challenge to accurately simulate weather phenomena. The current study quantifies the sensitivity of WRF simulation results to ~~physical parameterizations and numerical setup, and aims at identifying most physics parameterizations and model setup to identify best~~ suitable combinations for modeling ~~the storm Ciara EWE passing over 3 EWE detected from SCADA data collected at~~ the Belgian offshore wind farms ~~in the North Sea in February 2020. A simplified~~. This multi-variant ~~sensitivity analysis considering multi-event sensitivity analysis considers~~ 12 combinations of ~~PBL, cumulus, microphysics, temporal resolution of lateral boundary conditions and geospatial nested domain configurations~~ are investigated physics combinations comprising 3 PBL schemes, 4 cumulus schemes, 3 microphysics schemes, and hourly

versus 3-hourly update intervals of LBC. The remainder of this ~~paper-article~~ is structured as follows. Firstly, ~~the storm Ciara extreme weather event~~ a description of EWE is introduced in Section 2.1.2. Next, the numerical methodology and modeling setup are introduced in Section 3, where also the design of the sensitivity matrix is further described. Subsequently, the results are presented and discussed in Section Sect. 3. The simulation results and discussions are presented in Sect. 4. Lastly, conclusions and ~~perspective are exposed in Section~~ future prospects are presented in Sect. 5.

2 Description of the ~~storm Ciara event~~ events

The ~~case study selected in this study is~~ selection of the events in this study is motivated by the occurrence of fast changes in wind direction accompanied by severe yaw misalignment leading to significant power loss as observed by a Belgian offshore wind farm in the North Sea. The methodology utilized to identify these events modifies the approach defined by Hannesdóttir and Kelly (2019) to include yaw misalignment. The wavelet analysis considers a minimum threshold to identify anomalous changes in wind direction accompanied by severe yaw misalignment experienced by several wind turbines. Severe yaw misalignment potentially has adverse effects on the operational lifetime and fatigue loading of a wind turbine (Wan et al., 2015; Bakhshi and Sandborn, 2016; Laino and Hansen, 1998; Damiani et al., 2018), highlighting its importance and relevance in this study. The SCADA analysis for the identification of these events includes confidential error codes and data that are protected under a non-disclosure agreement, therefore no further details can be provided herein.

Three case studies are considered in this sensitivity analysis, namely, Storm Ciara on 10 February 2020, a low-pressure system on 24 December 2020, and a trough passage on 27 June 2020. The radar data presented therein is not publicly available, but was retrieved through a bilateral agreement with the Royal Meteorological Institute of Belgium (RMI-B). A brief synopsis of these events is presented in the following sub-sections.

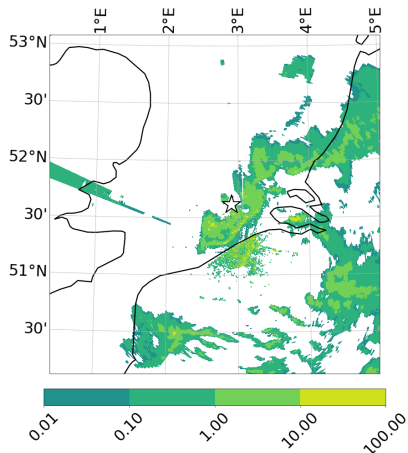
2.1 Case study 1: Storm Ciara

~~Storm Ciara is~~ one of the first extratropical cyclones to hit the European continent in the year 2020, ~~storm Ciara. Originating on the Atlantic Ocean, the storm occurred on occurring on~~ 10 February 2020 over the Belgian North Sea, ~~transpiring from North America.~~ Storm Ciara originated in the Atlantic Ocean, moving from the North American continent (starting 3 February 2020) to the European continent (16 February 2020). Storm Ciara swept across the majority of western Europe including the United Kingdom and Norway, bringing in heavy precipitation and strong winds with a maximum recorded wind gust of 219 km h⁻¹ at Cap Corse, Corsica, France¹. Over Belgium, the ~~Royal Meteorological Institute—Belgium (RMI-B)~~² reported wind gusts of up to 115 km h⁻¹ in Ostend, located at the Belgian offshore coast, with precipitation averaging 28 mm in few hours heavy precipitation accompanied by strong winds and thunderstorms over the local region.

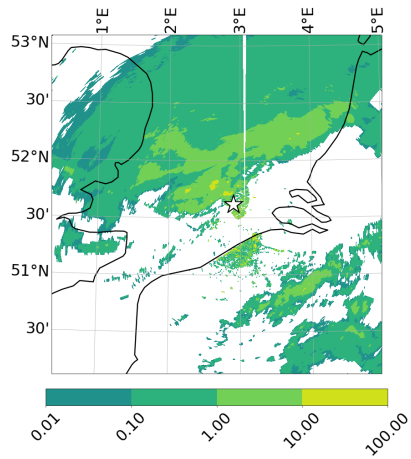
The selection of this case is motivated by the occurrence of fast changes in wind direction as observed by offshore wind farms located in the Belgian offshore concession zone at several moments during the storm.

¹<https://www.meteo-paris.com/actualites/retro-meteo-2020-les-evenements-climatiques-marquants-en-france>, website consulted on 21 April 2022.

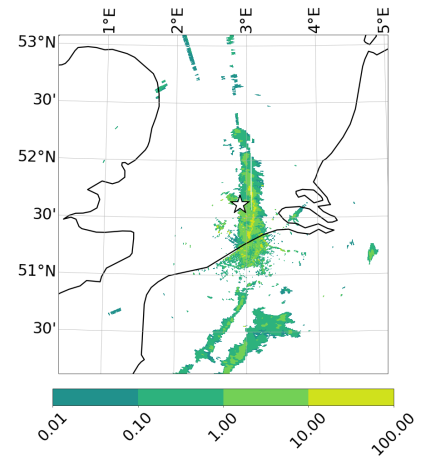
²<https://www.meteo.be/nl/info/nieuwsoverzicht/storm-ciara>, website consulted on 21 April 2022.



(a) Observed RADAR reflectivity on 10 February 2020 at 04:00 provided by a C-band Doppler RADAR located near the Belgian offshore coast in Jabbeke, Belgium.

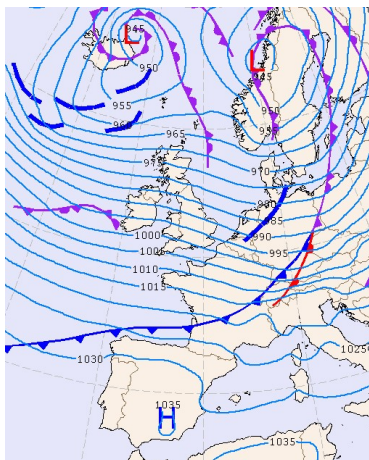


(b)

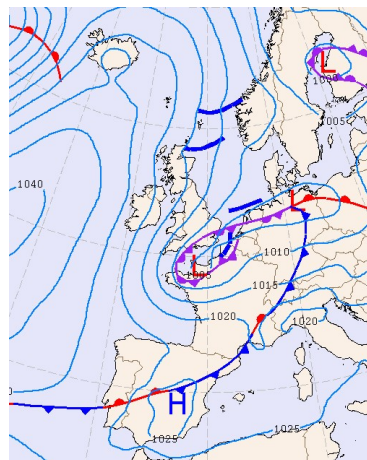


(c)

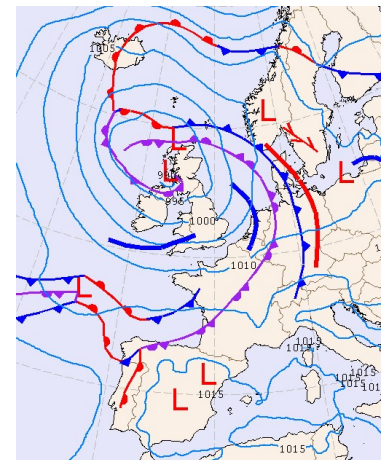
Figure 1. Observed precipitation rate in mm h^{-1} provided by a C-band Doppler radar located in Jabbeke on the Belgian coast. The star in the plots represents the offshore wind farm of interest. For the meteorological events: (a) Storm Ciara on 10 February 2020 at 04:40 UTC. (b) Low-pressure system on 24 December 2020 at 02:00 UTC. (c) Trough passage on 27 June 2020 at 15:30 UTC.



(a)



(b)



(c)

Figure 2. Synoptic maps provided by RNMI. (a) Storm Ciara on 10 February 2020 at 06:00 UTC. (b) Low-pressure system on 24 December 2020 at 00:00 UTC. (c) Trough passage on 27 June 2020 at 18:00 UTC.

During the early hours of ~~storm-Storm~~ Ciara on 10 February 2020, ~~a-commercial-an~~ offshore wind farm ~~observed-recorded~~ fast changes in wind direction of -40° ~~over few minutes accompanied by concentrated rainfall over a short period of time. This event was investigated with RADAR observations provided by~~ accompanied by severe yaw misalignment and concentrated rainfall. An RMI-B ~~for brevity the time-stamp radar snapshot~~ at 04:40 UTC is presented in Fig. ~~??1a~~, illustrating the presence of a bow-echo ~~transpiring-moving~~ from the British isles to Belgium, ~~indicative-an indication~~ of a possible micro-burst phenomena (Fujita, 1978). ~~Fast-phenomenon~~ (Fujita, 1978). Synoptic maps by the Royal Netherlands Meteorological Institute (RNMI³) presented in Fig. 2a indicate a trough passage during this period. Further, precipitation data from a wind profiler located within the wind farm highlights fast changes in wind direction ~~are potentially influential of the state of power and grid balances and have been found to be potentially harmful for operational conditions and lifetime of wind turbines as demonstrated~~ in ~~accompanied by sudden precipitation during the period of interest at 04:40 UTC, presented in Fig. 3.~~

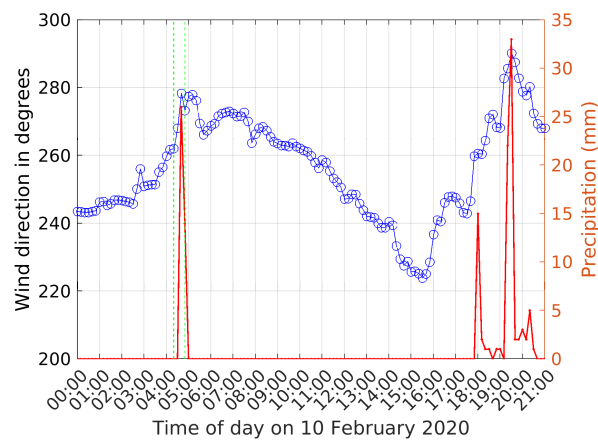


Figure 3. Precipitation observed by the offshore wind farm plotted against 10-min averaged wind direction (SCADA data). The highlighted the period of interest, in green, at 04:40 UTC is observed to accompany sudden precipitation.

2.2 Case study 2: Low-pressure system

On 24 December 2020, the Belgian offshore wind farms observed heavy precipitation accompanied by fast changes in wind direction. Synoptic maps presented in Fig. 2b indicate the presence of a low-pressure system over the North Sea. Radar observations from the RMI-B indicate large precipitation cells over the Belgian North Sea, presented in Fig. 1b. SCADA data records fast changes in wind direction of 100° at 02:00 UTC accompanied by severe yaw misalignment and precipitation.

2.3 Case study 3: Trough passage

³<https://www.knmi.nl/nederland-nu/klimatologie/daggegevens/weerkaarten>, website consulted on 21 April 2022.

On 27 June 2020, the studies by Damiani et al. (2018); Bakhshi and Sandborn (2016). With the addition to RADAR data and the availability of operational wind farm data (SCADA), the premise of this study provides a unique opportunity to investigate storm Ciara as felt by the Belgian offshore wind farms and formulate an informed decision on the optimum WRF set-up in the context of wind energy applications. Experienced fast changes in wind direction and sudden precipitation during the afternoon hours around 15:30 UTC. The synoptic maps provided by the RNMI indicate the presence of a low-pressure system over the British Isles, with a trough passage across the Belgian North Sea, presented in Fig. 2c. Radar observation provided by the RMI-B indicate the presence of precipitation cells over the offshore wind farms, presented in Fig. 1c. The operational SCADA data records fast changes in wind direction of 60° during this hour accompanied by severe yaw misalignment and precipitation.

3 Methodology and Model setup, methodology and performance metrics

This sensitivity study considers the WRF-ARW-WRF model version 4.2.2 to study the case of storm Ciara, evaluated against RADAR and SCADA observations from the Belgian North Sea. The model parameters simulate the case studies described in Sect. 2. The following sections describe the part of the WRF model setup common to all simulations, individual run setups used in the sensitivity study, and performance evaluation metrics for comparison to observational data. This evaluation uses operational wind farm SCADA data for its quantitative analysis of wind direction and wind speed. Additionally, radar data from RMI-B allows for a qualitative perspective on precipitation. By combining these observational datasets, the premise of this study provides a unique opportunity to investigate EWE as experienced by an offshore wind farm to determine suitable WRF setups in the specific context of wind energy applications.

3.1 Common model setup

The common model parameters considered for all WRF simulations are summarized in Table 1. The baseline horizontal resolution in the largest grid spacing of the parent domain d01 is 27 km, while the 1-way nested domains are sequentially refined with by a factor of 3, resulting in resolutions horizontal grid spacings of 9, 3, and 1 km for d02, d03, and d04 respectively. In the vertical direction, 57 terrain following pressure levels are considered with a model top pressure at 1000 Pa. The vertical velocity damping option based on the Courant-Friedrichs-Lewy condition as implemented in WRF is also turned on. The model is initialized on 9 February 2020 at 00:00, followed by a A time step of 20 s is considered for parent domain d01, while the time step of nested domains is sequentially refined by a factor of 3. The initial and LBC are derived from ERA5 reanalysis (Hersbach et al., 2020). WRF simulations were initialized with a spin-up period of 24 hours. Subsequently, the model is run for all case studies. In addition, an evaluation period of 21 hours from 00:00 to 21:00 UTC on 10 February 2020 is considered for Storm Ciara. For the low-pressure system and trough passage, an evaluation period of 6 h from 00:00 to 21:06:00, which adequately captures the storm event on the Belgian North Sea. The long-wave and short-wave radiation physics schemes are kept constant as UTC on 24 December 2020 and 6 h from 12:00 to 18:00 UTC on 27 June 2020 are considered, respectively. The simulations have been performed as a continuous run including spin-up and evaluation periods.

275 Therein, the selected evaluation periods adequately capture the time periods of interest for respective case studies as described in Sect. 2. The one-way nested domain configuration common to all simulations in this study is presented in Fig. 4.

The Rapid Radiative Transfer Model (RRTMG) (Iacono et al., 2008) for longwave and shortwave radiation physics is used by all simulations. Similarly, the land-surface interactions are kept constant as land-surface interactions are defined by the unified Noah land surface model (Tewari et al., 2004). The PBL, cumulus, and microphysics schemes are varied amongst the
 280 mentioned options as described in Table 1.

Table 1. WRF model setup and common parameters for all simulation runs. Parameters ~~The varied in this sensitivity study~~ physics parameters are ~~shown highlighted~~ in italics. Scale-aware physics ~~parametrizations~~ parameterizations are underlined.

Numerical setup	
Case studied <u>Nested domains (1-way nesting)</u>	storm Ciara, 10 February 2020 <u>Nested domains 4</u>
Horizontal resolution <u>grid spacing</u>	27 km (d01) × 9 km (d02) × 3 km (d03) × 1 km (d04)
Terrain following vertical levels	57
Model top pressure	1000 Pa
Time step <u>Time-steps for domain configuration</u>	20 s (d01), <u>6.67 s (d02), 2.22 s (d03), 0.74 s (d04)</u>
Spin-up period	24 h (9 Feb 2020 00:00 – 10 Feb 2020 00:00) ERA5 reanalysis
Simulation time <u>21 h (10 Feb 2020 00:00 – 10 Feb 2020 21:00) Initial</u>	
<u>Lateral & boundary conditions</u>	
Domain size <u>Evaluation time, additional to spin up</u>	<u>type 1 / type 2 (Fig. 4) 21 h (Storm Ciara) and 6 h (Low-pressure system and trough passage)</u>
<u>Boundary update frequency</u> interval	<u>1h / 3h</u>
Physics parameterizations	
Radiation	RRTMG <u>radiative</u>
Land surface	unified Noah land-surface
<i>PBL</i>	<i>MYNN / Shin-Hong / Zhang</i>
<i>Microphysics</i>	<i>WSM5 / Thompson / Morrison</i>
<i>Cumulus</i>	<i>KF / GD-3D / msKF / GF</i>

3.2 Individual run setups

In order to sufficiently categorize and distinguish the key features of different WRF physics parameterizations and options available ~~in WRF~~, a combination of different simulation pairs ~~in the multi-variant sensitivity~~ Table 2 is as described in Table 2 are considered. A total of 12 WRF simulations are categorized into different simulation pairs (A – ~~KJ~~) assigned to ~~either~~
 285 ~~variations of the temporal resolution of lateral boundary conditions, cumulus, microphysics, PBL schemes or geospatial domain configuration~~ variations of update interval of LBC, PBL, cumulus, and microphysics schemes. For each ~~variation of the varied~~

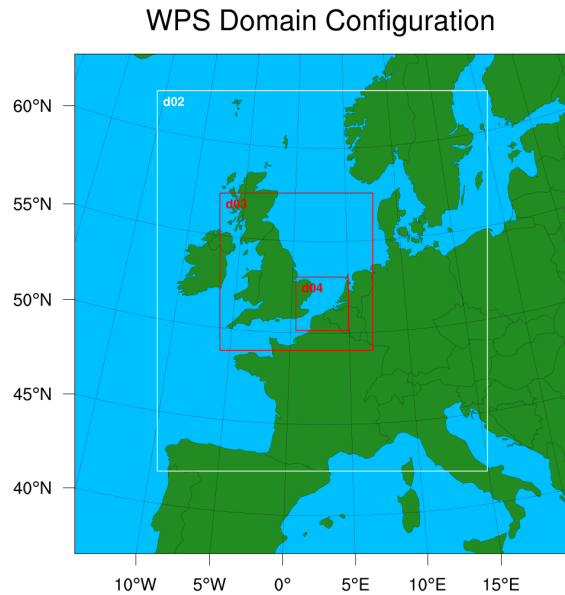


Figure 4. WRF and WRF Post-processing System (WPS) nested domain configuration (1-way nesting) considered common to all simulation runs in this study.

parameters, at least 2 different simulation pairs ~~have been performed~~ are considered. For example, simulation pairs A and B are assigned to the variation in ~~temporal resolution of the lateral boundary conditions, as within each pair only this temporal resolution is modified.~~ update interval of the LBC. More specifically, the simulation pairs considered are as follows. The

290 sensitivity to ~~1-hourly versus hourly and~~ 3-hourly ~~temporal resolution of lateral ERA5 boundary conditions is assessed with~~ update intervals of LBC are assessed in simulation pairs A and B. Further, the sensitivity to scale-aware ~~and (SH and Zhang)~~ and non-scale-aware ~~cumulus parameterizations is evaluated through (MYNN) PBL schemes is evaluated in~~ pairs C, D, and E. The ~~considered cumulus schemes consist of~~ sensitivity to scale-aware ~~GF, msKF, and (msKF and GF)~~ and non-scale-aware (KF and GD-3D ~~cumulus schemes~~) cumulus parameterizations is evaluated in pairs F, G, and H. Given the convection-permitting

295 resolutions of 3 km and 1 km for d03 and d04 respectively, the non-scale-aware KF model is explicitly turned off in these domains in simulation ~~eases 1-7~~ runs 2, 3, 5, and 6⁴. For the scale-aware cumulus models, this explicit deactivation is omitted, as they were specifically designed for operation on the verge of convection-permitting resolutions (Grell and Freitas, 2014; Zheng et al., 2016; Huang et al., 2020). The impact of microphysics schemes WSM5, Thompson, and Morrison is illustrated through pairs ~~F and G~~. Next, the comparison between scale-aware PBL ~~Shin-Hong I~~ and non-scale-aware PBL ~~MYNN is~~ shown in pairs H and I. Finally, pairs J and K determine sensitivity to the geographical domain size, types 1 and 2 as defined

300 ~~in Fig. 4, where the overall domain sizes are increased in type 2, including an increased northeasterly upstream fetch towards~~

⁴It was verified that this approach ~~results resulted~~ in better reproduction of precipitation cells and lower error metrics ~~as discussed below~~

the zone of interest in the Belgian North Sea (central in domain d04). Each simulation pair justifies to serve as independent sets of simulations to judge the influence of a varied WRF parameter.

Table 2. WRF simulation runs and respective simulation pairs considered for each of the varied parameter in this sensitivity analysis.

Simulation run#	ERA5 LBC updates	PBL scheme	Cumulus scheme	Microphysics scheme	Update interval pairs	PBL pairs	Cumulus pairs	Microphysics pairs
1	3 h	MYNN	msKF	WSM5	A			
2	3 h	SH	KF	WSM5	B			
3	1 h	MYNN	KF	Thompson		C		
4	1 h	MYNN	msKF	WSM5	A	D		
5	1 h	SH	KF	WSM5	B		F	I
6	1 h	SH	KF	Thompson		C	G	
7	1 h	SH	msKF	WSM5		D	F	J
8	1 h	SH	msKF	Thompson			G	
9	1 h	SH	msKF	Morrison			H	
10	1h	SH	GD-3D	Morrison				
11	1 h	SH	GF	Morrison		E		
12	1 h	Zhang	GF	Morrison				
13	Ensemble average							

Domain configuration 1 Domain configuration 2 The WRF nested domain configurations considered in this study. (a) Baseline domain configuration, domain 1 (b) Extended domain configuration 2, domain 2.

3.3 Performance metrics and observations

The simulated wind direction and wind speed from WRF runs horizontal wind speed are evaluated against front-row 10 minute averaged SCADA data from the commercial southwestern front row of an offshore wind farm located in the Belgian North Sea. The precipitation rate using RADAR reflectivity is evaluated against RADAR data from RMI-B. Further, RADAR reflectivity data is qualitatively compared to WRF-simulated reflectivity for 04:40 on 10 February 2020. Model accuracy is assessed using a standard Mean Absolute Error (MAE), as well as the Kantorovich distance as described in Wang and Basu (2016). The Kantorovich distance d , initially formulated in Kantorovitch (1958), is defined as the solution to the optimal transport problem transforming a discrete source signal $a_i (i = 1 \dots m)$ into a target signal $b_k (k = 1 \dots n)$,

$$d = \min_{x_{i,k}} \sum_{i=1}^m \sum_{k=1}^n r_{i,k} x_{i,k},$$

where $r_{i,k}$ is a cost associated with the phase shift between indices i and k , and $x_{i,k}$ is related to the difference in amplitude between a_i and b_k . In this regard, the Kantorovich distance d is a metric for the similarity between two signals, in this case

~~WRF results and field data, detecting both amplitudes and temporal phase shifts, hence supplementing the standard MAE as a point metric only considering pairwise differences in amplitudes at the same time instance.~~

for wind direction and wind speed. To recover a single performance metric, MAE and Kantorovich of wind direction and
320 wind speed are normalized to the so-called ~~N~~normalized Normalized Euclidean Distance (NED), given by $NED = \sqrt{MAE_N^2 + d_N^2}$,

$NED = \sqrt{MAE_{WDn}^2 + MAE_{WSn}^2}$. NED is defined as the resultant of normalized mean absolute error MAE_N , and normalized
~~Kantorovich distance~~ d_N of wind direction MAE_{WDn} and normalized mean absolute error of horizontal wind speed MAE_{WSn}
for all simulation runs. Normalization is performed with the mean over all simulations. ~~This study considers a univariate
analysis for wind variables and precipitation using performance metrics, NED and Kantorovich distance, respectively~~

325 Precipitation fields are qualitatively compared between WRF simulations. The simulated radar reflectivity is ~~extracted for a
single point in space at the location of the offshore wind farm and converted to rainfall~~ converted to precipitation rate using the
Marshall and Palmer equation (Marshall and Palmer, 1948) and is evaluated against reflectivity data from the dual-polarization
~~C-band radar from relation (Marshall and Palmer, 1948).~~

This study also evaluates the performance of an ensemble average compared to single deterministic simulation runs. The
330 ensemble average is defined as the mean of all simulation runs considered for a given case study. In this study, ensemble
members are initialized with identical initial conditions from ERA5 reanalysis. Subsequently, variability in the RMI-B, located
~~in Jabbeke at the Belgian North Sea coast.~~ ensemble average is only caused by the variation in update interval of LBC and
physics parameterizations. Therein, the current definition of ensemble average differs from traditional ensemble forecasts,
where variations in initial conditions are also considered, see, e.g., Wilks (2019).

335 4 Results and Discussion

The ~~summary of evaluated metrics for all WRF simulations is presented in Table ??.~~ The table presents MAE, Kantorovich
~~distance and NED for the complete simulation list conducted in this~~ following sections 4.1, 4.2, and 4.3 present results
and discussions for the overall trend in simulation results evaluated against SCADA data for Storm Ciara, the low-pressure
system, and the trough passage cases, respectively. The MAE and NED are presented in performance evaluation tables under
340 each case study. The table sequence is organized in the order of increasing ~~computational costs and complexity of physics
parameterizations~~ complexity of the combination of physics parameterizations and shorter update interval, starting with low
~~resolution of lateral boundary conditions,~~ the longer update interval of LBC and non-scale-aware physics parameterizations,
to scale-aware physics parameterizations with ~~higher resolutions of lateral boundary conditions.~~ Cell colors are assigned to
~~indicate the better metric-specific value in green per column.~~ The average NED is calculated as the average of horizontal the
345 shorter update intervals of LBC. Cells are colored based on a set of 5 categories between red and green. Categories are defined
to cover 20% of the range between smallest and largest values for the considered metric. In this way, results are categorized into
best (green, with errors in the 20% lowest range), good (light green), average (yellow), poor (light red), and worst (dark red). In
addition to this high-level assessment of setups, sections 4.4, 4.5, and 4.6 discuss simulation pairs addressing the influence of
specific combinations of physics parameterizations and update interval of LBC. The simulated wind direction and wind speed

350 ~~NED are quantitatively evaluated against SCADA data. Finally, Sect. 4.7 provides a synthesis of the observations in this study.~~

~~Before discussing the influence of specific parameterizations, we discuss some general trends~~

4.1 Case: Storm Ciara

The MAE of wind direction, horizontal wind speed, and the NED for the Storm Ciara runs is presented in Table 3. Overall, an increasing conformity with observations is observed for simulations with scale-aware physics parameterizations coupled with a larger domain configuration and higher resolution of lateral boundary conditions. The for simulation runs 2 through 12, relatively lower MAE values for wind direction and wind speed are observed, with a maximum of 9.26° and 2.72 m s^{-1} respectively. Using NED as the evaluation metric, the best-case setup in this study is determined to be simulation case 12 with the lowest average NED. Case 12 encompasses run 7, with a NED value very similar to the ensemble average (< 1 % difference). Run 7 uses the scale-aware Shin-Hong SH PBL scheme coupled with a double moment 6-class Morrison the scale-aware msKF cumulus scheme, single moment 5-class WSM5 microphysics scheme, a scale-aware GF cumulus scheme and hourly ERA5 1h reanalysis dataset as the lateral boundary conditions. The lateral boundary condition updates. In a general sense, simulation runs 7 through 10 observe the lowest overall NED. These simulation runs consider scale-aware PBL and cumulus parameterizations coupled with hourly update intervals of LBC. A qualitative analysis of wind direction and wind speed comparisons for the complete list of WRF simulations along with their ensemble average timeseries for all simulation runs highlighting the ensemble average and the best-case setup is presented in Fig. ?? . Further, Fig. ?? illustrates individual and ensemble precipitation results as compared to RADAR data. Qualitatively, the simulations capture the 5. Compared to the SCADA reference data, the changes in wind direction reasonably well when compared to SCADA data. On the other hand, matching wind speeds and precipitation seems significantly are captured reasonably well by all runs, with the ensemble average capturing the general transience of wind direction better than the best-case setup. However, accurately capturing the variability on wind speed is found to be more challenging, as shown by the large spread among different modeling setups in the afternoon and evening hours. Interestingly, despite the average NED metric not indicating the ensemble average as a clear winner amongst the simulation cases (due to relatively high errors on wind directions), the precipitation accuracy is greatly improved over the best case 12.

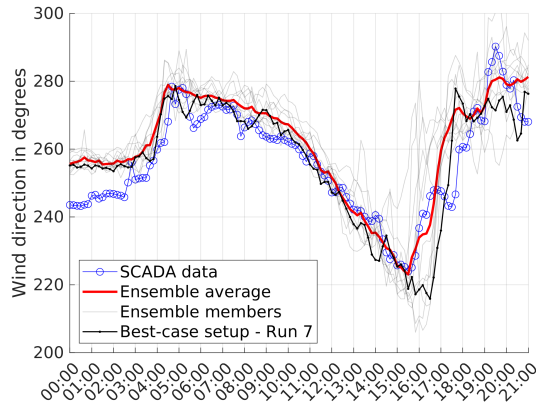
375 ~~Ensemble average of WRF RADAR reflectivity evaluated against observational RADAR data from RMI-B. Ensemble members shown in gray.~~

4.2 Case: Low-pressure system

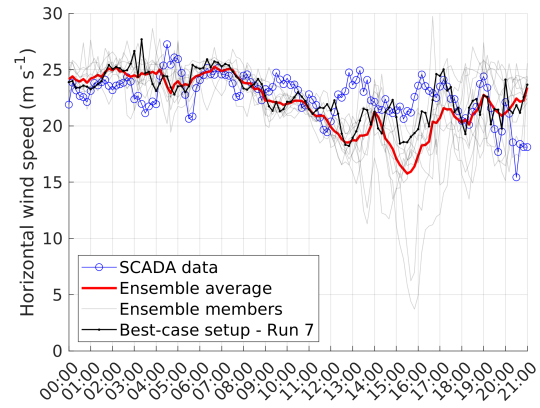
The remainder of this section focuses on the sensitivity to individual parts of the modeling chain. Firstly, the domain configuration is discussed in Sect. ?? . Next, the influence of the PBL schemes is shown in Sect. ?? , followed by the temporal resolution of initial boundary conditions in Sect. 4.4. Subsequently, Sect. 4.6 elaborates on cumulus and microphysics schemes. Finally, Sect. 4.7 provides a synthesis of the observations in this section.

Table 3. Evaluation of wind velocity Performance metrics MAE and NED for wind direction and wind speed from all WRF simulations vs evaluated against SCADA data averaged over for the first row case of wind turbines using metrics mean absolute error, Kantorovich distance and mean normalized euclidean distance Storm Ciara. The best-case setup considering NED is simulation run 7. The minimum metric specific values are underlined.

Simulation run#	ERA LBC updates	PBL scheme	Cumulus scheme	Microphysics scheme	Wind direction MAE (degrees)	Wind speed MAE ($m s^{-1}$)	NED (-)
1	3 h	MYNN	msKF	WSM5	10.46	3.88	2.08
2	3 h	SH	KF	WSM5	8.48	2.57	1.51
3	1 h	MYNN	KF	Thompson	9.26	2.72	1.63
4	1 h	MYNN	msKF	WSM5	8.61	2.54	1.51
5	1 h	SH	KF	WSM5	7.68	2.47	1.41
6	1 h	SH	KF	Thompson	8.37	2.51	1.48
7	1 h	SH	msKF	WSM5	6.59	<u>1.78</u>	<u>1.11</u>
8	1 h	SH	msKF	Thompson	6.69	1.89	1.15
9	1 h	SH	msKF	Morrison	7.17	1.89	1.20
10	1 h	SH	GD-3D	Morrison	<u>5.59</u>	2.25	1.17
11	1 h	SH	GF	Morrison	7.17	2.67	1.43
12	1 h	Zhang	GF	Morrison	8.69	1.84	1.34
13	Ensemble average				5.88	2.04	1.12



(a)



(b)

Figure 5. Ensemble-average of WRF on Timeseries plots wind direction (a) and wind speed plotted along with the ensemble average and best-case setup simulation run 7 for the case of Storm Ciara. (a) Wind direction. (b) evaluated against SCADA data Wind speed. Ensemble members shown in gray.

4.3 Simulation pair: Domain configuration

The current section investigates the extent to which geographical domain size and upstream fetch affect the generation of fine-scale wind variations, precipitation cells and convective systems for the current case of storm Ciara. In literature, various studies have provided an equivocal literature on the influence of domain size on WRF simulations ranging from better representation of convective systems in simulations with larger domains and vice-versa (see e.g. Bhaskaran et al. 1996; Yu et al. 2021; Wang et al. 2021). Simulation pairs J and K represent four WRF simulations with two different domain configurations (with identical spatial resolution) as presented. summary of the performance evaluation metrics for the low-pressure system is presented in Table 4. The best-case setup is found to be simulation run 2, comprising scale-aware SH PBL coupled with non-scale-aware KF cumulus, WSM5 microphysics, and hourly ERA5 reanalysis data as LBC. Unlike the case of Storm Ciara, no trend in better results for simulation runs combining scale-aware PBL, scale-aware cumulus, and higher-order microphysics is found (i.e., simulation runs 6 through 10 in Table 4). The overall trend in wind speed MAE results shows simulations using the MYNN PBL scheme, i.e., runs 1, 3, and 4, to perform poorly. Similar to the case of Storm Ciara, the best-case results are found to be very similar to the ensemble average, with a relative difference in NED of 3.2 %. However, it must be noted that the ensemble average tends to damp out the fast changes in wind direction, as plotted in Fig. 4. Note that J furthermore applies a standard MYNN PBL scheme whereas K uses the scale-aware Shin-Hong PBL scheme. Quantitative performance metrics for these simulation pairs are shown in Table ??, whereas a qualitative view is presented in the form of a wind direction time series and a snapshot of RADAR reflectivity in Figs. ?? and ?? respectively. 6 along with the best-case setup. A qualitative analysis on the timeseries indicates that all simulation runs capture the fast change in wind direction for the period of interest. However, these often exhibit a time lag compared to SCADA data.

Focusing firstly on the wind direction, it can be seen-

Table 4. Performance metrics MAE and NED for wind direction and wind speed from all WRF simulations evaluated against SCADA data for the low-pressure system case. The best-case setup considering NED is simulation run 2. The minimum metric specific values are underlined.

Simulation run#	ERA LBC updates	PBL scheme	Cumulus scheme	Microphysics scheme	Wind direction MAE (degrees)	Wind speed MAE (m s ⁻¹)	NED (-)
1	3 h	MYNN	msKF	WSM5	12.58	3.96	1.66
2	3 h	SH	KF	WSM5	<u>10.43</u>	<u>1.77</u>	<u>0.94</u>
3	1 h	MYNN	KF	Thompson	13.17	4.28	1.78
4	1 h	MYNN	msKF	WSM5	12.47	4.40	1.79
5	1 h	SH	KF	WSM5	13.60	2.19	1.20
6	1 h	SH	KF	Thompson	21.92	1.95	1.62
7	1 h	SH	msKF	WSM5	16.92	2.16	1.37
8	1 h	SH	msKF	Thompson	15.74	2.30	1.34
9	1 h	SH	msKF	Morrison	19.81	3.12	1.73
10	1 h	SH	GD-3D	Morrison	15.97	2.32	1.35
11	1 h	SH	GF	Morrison	12.11	2.64	1.25
12	1 h	Zhang	GF	Morrison	15.54	2.15	1.29
13	Ensemble average				10.65	1.85	0.97

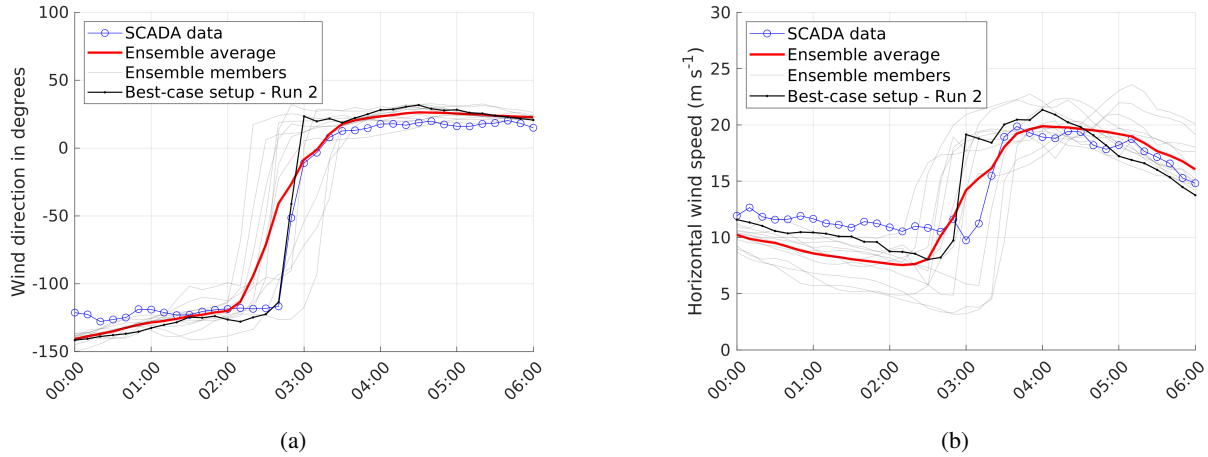


Figure 6. Timeseries plots wind direction and wind speed plotted along with the ensemble average and best-case setup simulation run 2 for the low-pressure system case. (a) Wind direction. (b) Wind speed.

4.3 Case: Trough passage

The performance evaluation metrics for the trough passage are presented in Table 5. Overall, the considered WRF setups are found to be highly sensitive to the combinations and type of physics parameterizations. Overall, MAE values are found to be higher than for the other two cases, indicating that accurately predicting wind direction and wind speed is more challenging for this particular event. No clear trend in any combinations of the sequence of simulation runs is found, in contrast to Storm Ciara and the low-pressure system. Interestingly, considering NED as the evaluation metric, the best-case setup is observed to be run 12 by a significant margin. Run 12 uses the scale-aware Zhang 3D PBL scheme, the scale-aware GF cumulus scheme coupled with Morrison microphysics and hourly ERA5 LBC updates. Simulation timeseries including the ensemble average and best-case setup are presented in Fig. ?? that the simulations with the larger spatial extent on Domain 2 (yellow lines) produce more short-timescale fluctuations, similar to fluctuations observed in the SCADA data, than those simulations performed on the smaller Domain 7. Qualitatively, simulation runs 1 configuration (orange lines). However, assessing the performance metrics in Table ??, this does not immediately translate to a clear reduction in NED for both pair J and K. Even though the Domain 2 fluctuations produce qualitative similarities with the SCADA data, their respective phase differences lead to large errors in MAE, which are only partially corrected for by the improved Kantorovich distance, thus resulting in an inconclusive trend regarding NED. Similar arguments can be made for the wind speed: even though the Domain 2 runs have the potential to provide higher levels of detail, through 11 underpredict the wind direction for the evaluation period, whereas wind speeds are underpredicted by all runs. Due to the match with SCADA data does not consistently improve based on our current metrics, joint poor performance and persistent offsets compared to SCADA data for all simulations except run 12, the ensemble average does not yield any better match to the data.

420 Turning to precipitation, Kantorovich distances from Table ?? clearly indicate simulations on larger Domain 2 to produce better results. Similar findings were also recently observed by Yu et al. (2021). Precipitation levels from WRF-simulated RADAR reflectivity versus observed RADAR reflectivity at 04:40 for simulation pair K are presented in Fig. ?. It is shown that whereas the Domain 2 WRF run indicates the presence of bow echo similar to the observed RADAR, little to no RADAR reflectivity is observed for simulations on Domain 1. For simulation pair J (not further shown here) domain 1 captured significantly less variation in precipitation fronts in comparison to domain 2. Overall, simulations on Domain 2 predicted better convective structures and precipitation cells that are found to be more representative of RADAR observations in comparison to the smaller Domain 1.

Evaluation of simulation pairs J and K, subject to two different domain configurations using MAE and Kantorovich distance.

Table 5. Performance metrics MAE and NED for wind direction and wind speed from all WRF simulations evaluated against SCADA data for the case of the trough passage. The best-case setup considering NED is simulation run 12. The minimum metric specific values are underlined.

Simulation run#	ERA LBC updates	PBL scheme	Cumulus scheme	Microphysics scheme	Wind direction MAE (degrees)	Wind speed MAE (m s ⁻¹)	NED (-)
1	3 h	MYNN	msKF	WSM5	15.12	4.22	1.39
2	3 h	SH	KF	WSM5	19.27	5.27	1.75
3	1 h	MYNN	KF	Thompson	16.04	3.47	1.32
4	1 h	MYNN	msKF	WSM5	11.64	4.46	1.28
5	1 h	SH	KF	WSM5	17.95	4.51	1.57
6	1 h	SH	KF	Thompson	19.84	4.60	1.68
7	1 h	SH	msKF	WSM5	15.05	5.31	1.57
8	1 h	SH	msKF	Thompson	16.34	5.02	1.58
9	1 h	SH	msKF	Morrison	10.97	3.78	1.13
10	1 h	SH	GD-3D	Morrison	17.73	4.57	1.57
11	1 h	SH	GF	Morrison	15.03	3.87	1.33
12	1 h	Zhang	GF	Morrison	<u>7.43</u>	<u>3.11</u>	<u>0.87</u>
13	Ensemble average				14.77	4.34	1.39

4.4 Update interval of lateral boundary conditions: Simulation pairs A and B

430 The effect of varying the update interval of ERA5 LBC between hourly and 3-hourly is investigated in this section. Simulation pairs A and B represent 4 WRF setups for each case study. Figure 8 shows the results for simulation pair A. Errorbars indicate one standard error of the sample mean. Starting with wind direction (Fig. 8a), hourly update intervals of ERA5 LBC are observed to perform better for Storm Ciara and the trough passage. For the low-pressure system, lower MAE is observed for hourly reanalysis data however, these values lie well within the standard error bars, therefore leading to inconclusive overall observations. For wind speed (Fig. 8b), a significant improvement on MAE is observed when hourly reanalysis data is used for Storm Ciara, with a 34 % reduction compared to 3-hourly data. In contrast, for the low-pressure system and the trough passage,

435

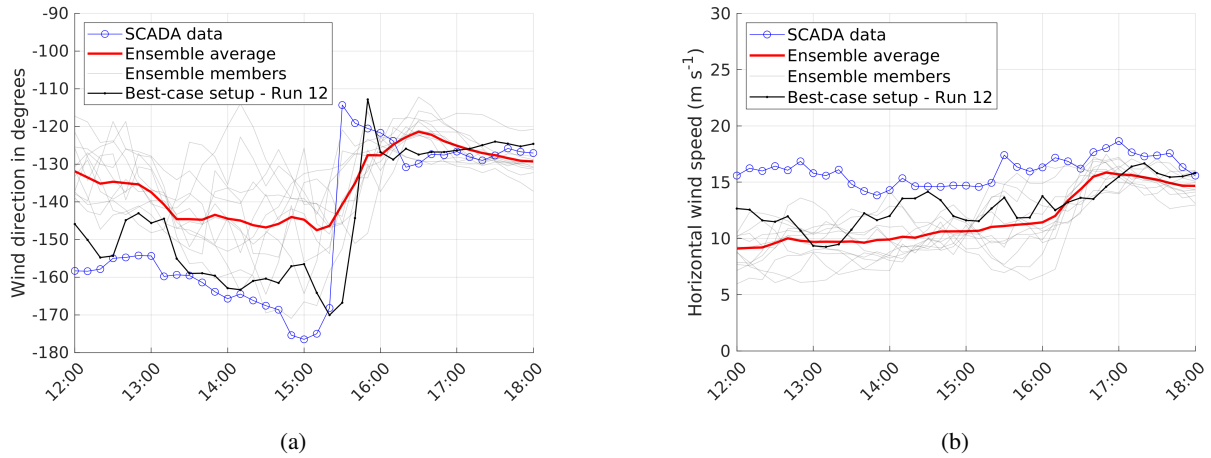


Figure 7. Timeseries plots wind direction and wind speed plotted along with the ensemble average and best-case setup simulation run 2 for the case of trough passage. (a) Wind direction. (b) Wind speed.

3-hourly reanalysis data produces lower MAE, however these values lie within the standard error. Overall, for simulation pair A, a distinction in better performance for hourly reanalysis is observed for Storm Ciara, however no significant benefit is observed for the other two cases.

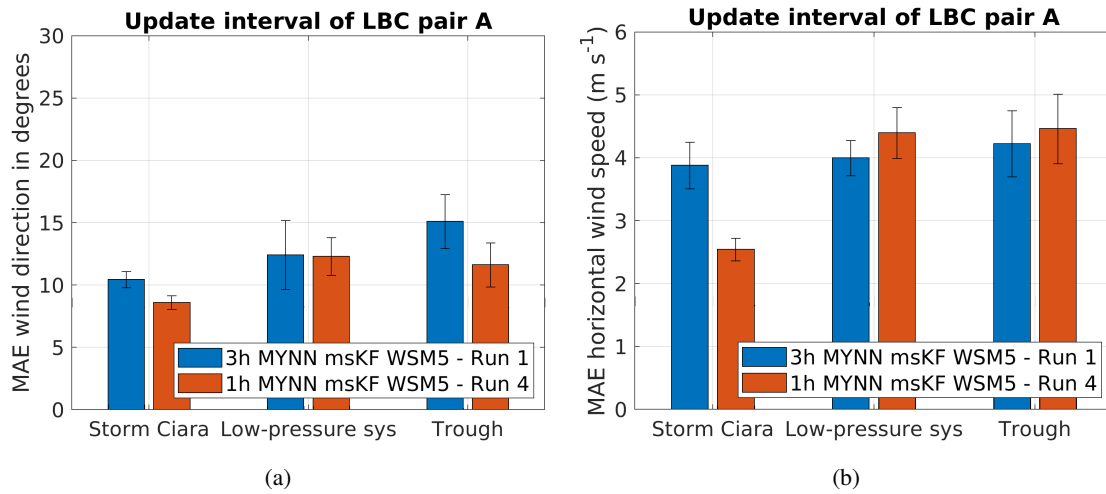


Figure 8. Performance evaluation for simulation pair A considering change in update interval of LBC, as described in Table 2. (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.

440 Similarly, Fig. 9 shows the results for simulation pair B. The MAE comparison for wind direction is presented in Fig. 9a, indicating statistically inconclusive results for all cases. Similarly, MAE results for wind speeds are presented in Fig. 9b.

indicating inconclusive results for Storm Ciara and the low-pressure system, yet a better performance with hourly reanalysis data in the case of the trough passage.

445 To summarize the overall inferences from both simulation pairs, hourly updates of LBC do not systematically lead to higher accuracy, although improvements are observed for certain combination of events and wind variables. Therefore, more frequent updates of LBC may prove advantageous when trying to capture certain fast transient weather events.

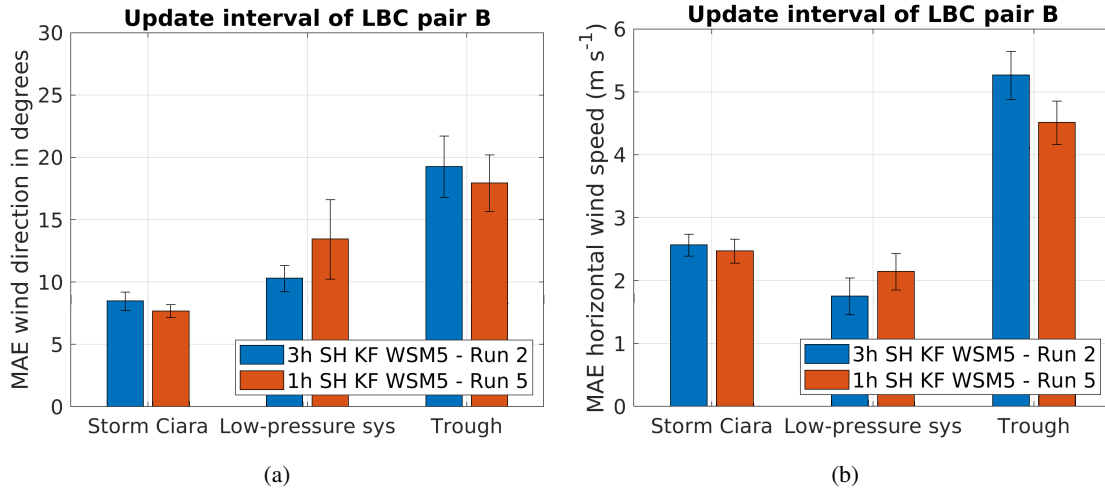


Figure 9. Wind direction time series at offshore wind farm location from SCADA data and WRF simulations. Performance evaluation for simulation pair B considering change in update interval of LBC, as described in Table 2. (a) Simulation pair J MAE comparison for wind direction. (b) Simulation pair K MAE comparison for wind speed.

Contours of RADAR precipitation (contour lines in foreground) and WRF precipitation (filled contours in background) for simulation pair K. Left: Simulation on smaller Domain 1. Right: Simulation on larger Domain 2. Color profile are scaled such that similar color patterns represent similar precipitation levels.

4.5 Simulation pair: Planetary boundary layer: Simulation pairs C, D, and E

450 In this section, the influence of using classical non-scale-aware PBL schemes versus using scale-aware PBL schemes is elaborated in simulation pairs C, D, and E. More specifically, the standard MYNN scheme is compared to the scale-aware Shin-Hong scheme by simulation pairs H (run on Domain 1, see previous section) and I (run on Domain 2). Note that these are the same simulations considered in the previous section, but compared in a different manner here. Table ?? contains quantitative performance metrics for these simulation pairs. Furthermore, 1D SH and scale-aware 3D Zhang PBL schemes. Simulation pairs C and D compare the influence of MYNN and SH PBL schemes and simulation pair E compares SH and Zhang PBL schemes.

455 Figure 10 presents the consolidated results for simulation pair I, the time series of wind direction and wind speeds are shown in Fig. 13 and a snapshot of RADAR reflectivity is presented in Fig. ??.

460 Considering wind direction and wind speed first, the simulations with the Shin-Hong PBL scheme observe better concurrence to SCADA data considering MAE. In some cases, the Kantorovich distance however contradicts MAE in favor of the MYNN PBL scheme, resulting in overall similar NED scores C. First, considering wind direction MAE (Fig. 10a), the SH PBL scheme performs better for the case of Storm Ciara. In contrast, for the low-pressure system and the trough passage, the MYNN PBL scheme shows better performance. Considering wind speed (Fig. 10b), no conclusive set of inferences are drawn for the case of Storm Ciara, as the lower MAE by SH lies within the range of the standard error. For the low-pressure system, 465 the SH PBL scheme outperforms the MYNN scheme in terms of wind speed, reversing the trend found for wind direction for Shin-Hong and MYNN. For wind speeds on the other hand, the resulting NED significantly favors scale-aware Shin-Hong. Finally, considering the averaged NED as the defining metric, better overall results were observed for simulations with the scale-aware Shin-Hong PBL scheme. For the trough passage, MYNN wind speeds outperform SH. Overall, for the simulation pair C, no clear conclusions can be drawn for Storm Ciara and the low-pressure system cases. In contrast, better performance 470 by the MYNN PBL scheme is observed for the trough passage.

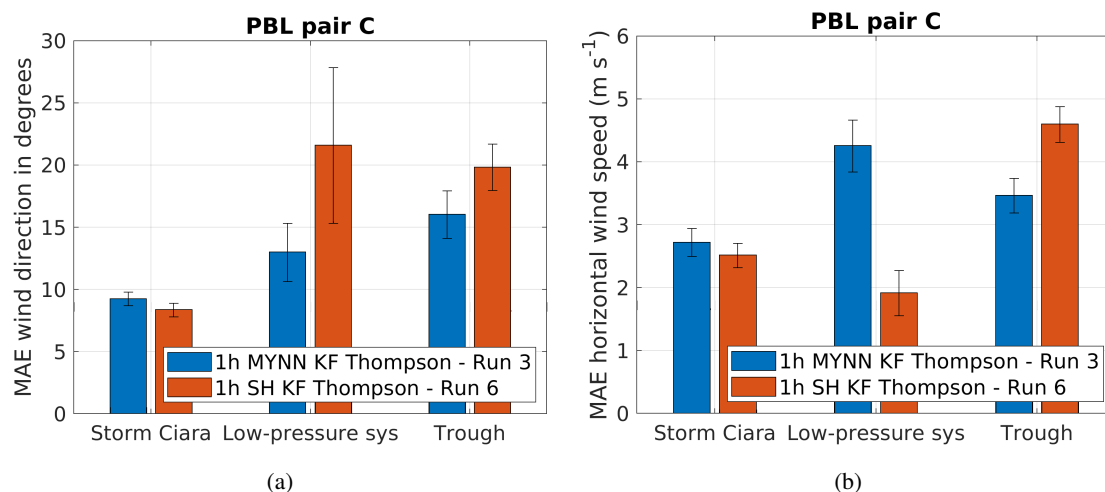
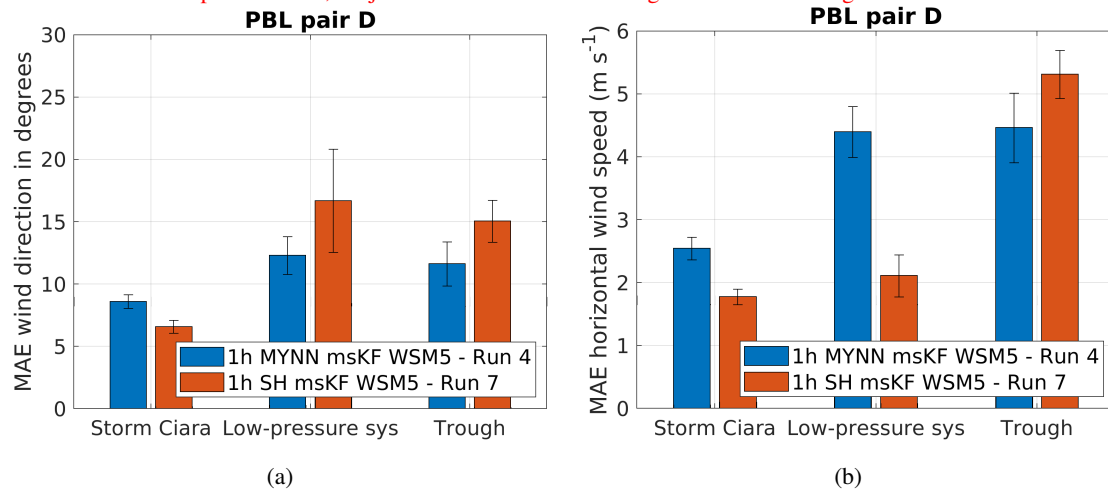


Figure 10. Performance evaluation for simulation pair C considering a change in PBL scheme, as described in Table 2. (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.

Even though we observed larger vertical velocity profiles and higher water vapour mixing ratios for Shin-Hong than for MYNN during times of high precipitation (see Fig. ??), the precipitation Kantorovich distance in Table ?? presents inconclusive results regarding which PBL scheme results in a closer match to RADAR data. The qualitative results as presented in Fig. ?? for simulation pair I indicate a better representation of the precipitation front for simulations with the Shin-Hong PBL scheme. Figure 11 illustrates MAEs for simulation pair D. Starting with wind direction (Fig. 11a), a clear advantage in utilizing SH PBL is observed for the case of Storm Ciara. However, this distinction is not observed for the low-pressure system. The MYNN PBL scheme performs better than the SH PBL scheme for the trough passage. This trend is found similar to simulation pair C. Considering wind speeds (Fig. 11b), SH outperforms MYNN by a significant margin for the case of Storm Ciara. A 475

similar distinction is observed for the low-pressure system. However, for simulation pair H (not further shown here), Shin-Hong scheme significantly underestimates the RADAR precipitation similar to pair C, for the trough passage this trend is reversed. Overall, current results show the scale-aware Shin-Hong PBL scheme performs generally better for wind variables in the current case, yet results for precipitation could indicate a dependency with cumulus and microphysics parameterizations, which has also been reported in literature (Hong and Dudhia, 2012; Choi and Han, 2020; Chen et al., 2021). for the simulation pairs C and D, a distinctly better performance by the SH PBL scheme is observed for the case of Storm Ciara. However, little to no conclusions can be drawn for the low-pressure system. For the trough passage, better performance is observed for MYNN PBL scheme.

Evaluation of simulation pairs H and I, subject to MYNN and Shin-Hong PBL schemes using MAE and Kantorovich distance.



Time series at offshore wind farm location from SCADA data and WRF simulations for simulation pair I. (a): Wind direction. (b): Wind speed.

Figure 11. Contours of RADAR precipitation (contour lines in foreground) and WRF precipitation (filled contours in background) Performance evaluation for simulation pair I. *Left:* MYNN-D considering a change in PBL scheme, as described in Table 2. *Right:* Shin-Hong PBL scheme (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.

To further distinguish between scale-aware PBL schemes of different complexity, Figure 12 presents results for simulation pair E, comparing the SH and the Zhang PBL schemes. The wind direction MAE results (Fig. 12a) show an advantage in using SH for the case of Storm Ciara. However, this distinction is not statistically significant for the low-pressure system. The Zhang PBL scheme outperforms the SH PBL scheme by a significant margin for the trough passage. Considering wind speed (Fig. 12b), the Zhang scheme outperforms the SH scheme for Storm Ciara and the trough passage. However, this trend is not statistically significant for the low-pressure system case. The trough passage observes better performance by the Zhang PBL scheme. Overall, for simulation pair E, no clear distinction in better performance for both wind direction and wind speed combined is observed for SH or Zhang PBL schemes for Storm Ciara and the low-pressure system cases. However, for the case of the trough passage, a clear advantage in using the Zhang PBL scheme is observed.

4.6 Simulation pair: Temporal resolution of lateral boundary conditions

The effect of varying temporal resolution or update frequency of the ERA5 lateral boundary conditions is investigated in this section. Simulation pairs A and B represent four WRF simulations in which the temporal ERA5 resolution is varied between hourly and three-hourly updates. Note further that A and B mutually differ in all setup parameters other than the cumulus scheme, which justifies them to serve as independent pairs to judge the influence of the temporal boundary condition resolution. A qualitative comparison of timeseries for the low-pressure system and trough passage cases is presented in Fig. 13b, indicating better performance by the Zhang PBL scheme to capture the transience in wind direction. However, this qualitative advantage is not observed in the case of Storm Ciara (not further plotted here).

Table ?? consolidates the results for both simulation pairs. The higher hourly temporal resolution for ERA5 lateral boundary conditions produced better results for wind direction, wind speed, and precipitation. Similar results were observed by Hamouda and Pasquero (2021) considering higher resolution of ERA5 dataset to simulate European extreme precipitation.

Evaluation of simulation pairs A and B, subject to 1h and 3h temporal resolution of lateral boundary conditions using MAE and Kantorovich

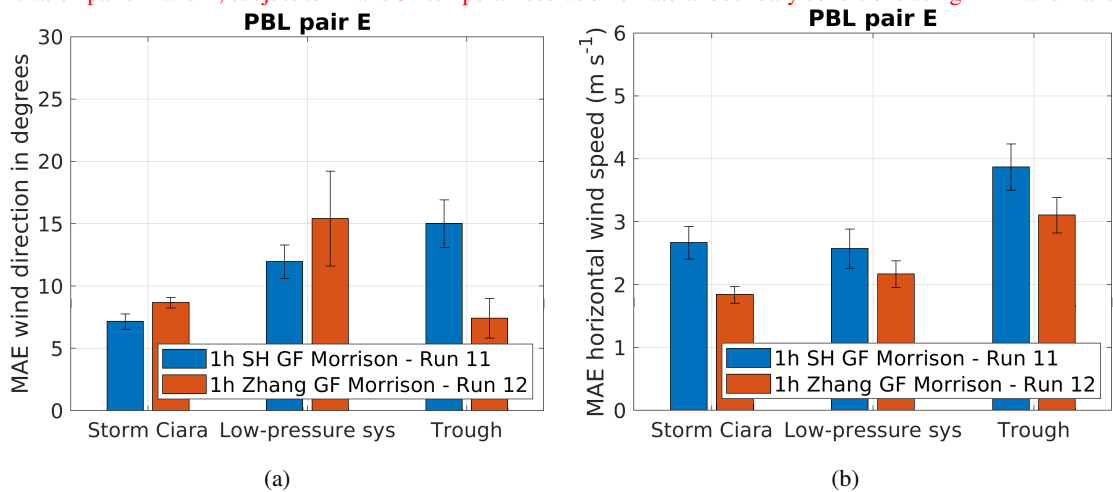


Figure 12. Performance evaluation for simulation pair E considering a change in PBL scheme, as described in Table 2. (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.

Figure ?? presents the time-series plots for wind direction for simulation pairs A and B. For both pairs, a slightly better qualitative match with SCADA is shown for the 1h resolution runs (orange lines) over the 3h runs (yellow lines). Specifically, in pair A, the afternoon through around 15:00 is better captured using the higher temporal resolution, whereas in pair B the peaks at 05:00 and 19:00 are modelled with higher accuracy. Note that the difference in terms of smoothness between pairs A and B can be attributed to the different domain configurations as discussed above in Section ?. A qualitative comparison of wind speeds revealed similar trends and is omitted here for brevity.

A qualitative look at the RADAR snapshot in Figure ?? for simulation pair B reveals that whereas the 1h resolution run reproduces the main frontal structure of the precipitation, the 3h resolution simulation does not show any reflectivity at all

515 The current results indicate a high sensitivity in wind direction and wind speed relative to the choice in PBL scheme. A single PBL scheme is not found to outperform the others for all 3 case studies. The simulation results indicate a possible dependency of PBL schemes to cumulus and microphysics parameterizations, which has also been reported in literature by Hong and Dudhia (2012), Choi and Han (2020), and Chen et al. (2021). For simulation pair A, which is not further plotted here, no apparent qualitative differences in RADAR reflectivity were observed.

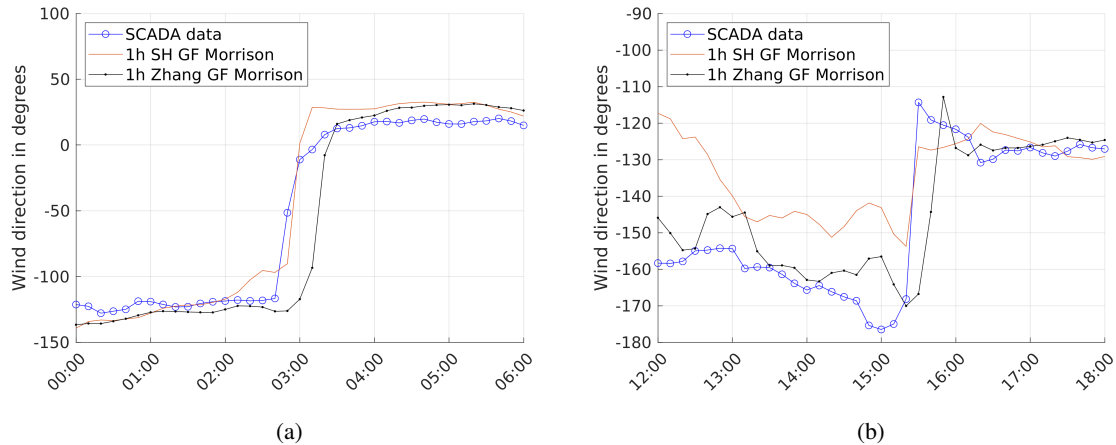


Figure 13. Time-series Comparison of wind direction at offshore wind farm location from SCADA data and WRF simulation time series considering a change in PBL scheme for simulation pair E. (a) Simulation pair B Wind direction plots for low-pressure system case. (b) Simulation pair A Wind direction plots for trough passage case.

520 Contours of RADAR precipitation (contour lines in foreground) and WRF precipitation (filled contours in background) for simulation pair B. Left: hourly boundary condition updates. Right: three-hourly boundary condition updates.

4.6 ~~Simulation pair:~~ Cumulus and Microphysics: Simulation pairs F, G, H, I, and J

This section presents ~~the consolidated and discusses the~~ results for cumulus ~~and microphysical scheme variation. Since both of these highlight a specific aspect of modeling simulation pairs F, G, and H, as well as microphysics simulation pairs I and~~
 525 J. As these two types of physics schemes both relate to the modeling of precipitation in WRF, they are considered jointly in this section. A set of 7 simulation configurations 8 WRF simulations is considered, covering a combination of 4 cumulus schemes (KF, GD-3D, msKF, and GF; the latter two being scale-aware) and 3 microphysical microphysics schemes (WSM5, Thompson, and Morrison; the latter two representing higher complexity in microphysical parameterization in increasing order of modeling complexity).

530 Firstly, Figure 14 depicts simulation results for pair F, which uses the SH PBL scheme coupled with the WSM5 microphysics parameterization while varying the cumulus scheme between KF and msKF. Figure 14 presents the results for MAE of wind direction and wind speed for all three test cases. Starting with wind direction (Fig. 14a), the wind results for cumulus parameterization pairs are discussed. Pairs C, D and E each allow to assess the influence of the cumulus scheme, while

considering one specific microphysics option, i. e. WSM5, Thompson, and Morrison for C, D, and E, respectively. Quantitative results msKF simulation produces better results for the case of Storm Ciara. However, for the low-pressure system case, KF produces lower MAE which lies within the standard error bars, disallowing statistically significant conclusions. Similarly, for the trough passage, the msKF scheme results in a negligible reduction in MAE. Therefore, conclusions can only be drawn in the case of Storm Ciara, indicating better performance with the msKF scheme on wind direction. Focusing on wind speed (Fig. 14b), again better results for msKF cumulus scheme are observed for the case of Storm Ciara. For the low-pressure system case both schemes produce similar MAE values. For the trough passage, KF performs better than msKF, reversing the trend found for wind direction. Summarizing, for simulation pair F, the case of Storm Ciara is more accurately predicted by the msKF cumulus scheme in comparison to the KF scheme. For the low-pressure system and trough passage cases, comparative results are inconclusive.

The results for simulation pair G are presented in Table ?? For Fig. 15. Pair G applies the SH PBL scheme coupled with the Thompson microphysics parameterization, while varying the cumulus scheme between KF and msKF. Overall, the current combination of physics schemes produce similar MAE on wind direction and wind speed, compared to simulation pair F. For Storm Ciara, msKF produces lower MAE in wind direction and wind speed. For the low-pressure system, no clear trend in performance is observed. For the trough passage, lower wind direction MAE is observed for msKF and lower wind speed MAE is observed for KF. Consolidating results for simulation pairs F and G, for the case of Storm Ciara, a clear improvement is observed when using the scale-aware cumulus parameterizations (msKF and GF) produce better overall NED in comparison to non-scale-aware schemes (KF and GD-3D). The overall outcome, as drawn from average NED, indicate best results for the scale-aware GF cumulus parameterization msKF cumulus scheme, however no conclusive statements can be made for the low-pressure system and trough passage cases.

Turning to the wind result sensitivity on microphysical parameterizations in Table ??, simulation pairs F (with KF cumulus) and G (with msKF cumulus) assess the impact of WSM5, Thompson and Morrison schemes. As shown in the table, these simulation pairs do not allow formulating a conclusive trend regarding the impact of microphysics on wind modeling in terms of average NED. However, considering both Table ?? and ??, an observation is that the impact of cumulus schemes on average NED is greater than that of microphysics, and that the best overall performance is obtained using Figure 16 shows results for simulation pair H, which applies the SH PBL scheme coupled with the Morrison microphysics, while varying the cumulus scheme between msKF, GD-3D, and GF. Starting with wind direction (Fig. 16a), Storm Ciara is much better captured by the GD-3D cumulus scheme. For the low-pressure system, no such trend in better performance is observed. For the trough passage, msKF outperforms both GD-3D and GF schemes. Considering wind speed (Fig. 16b), msKF performs better than GD-3D and GF for Storm Ciara. For the low-pressure system, GD-3D is found to be the best performer. For the trough passage, msKF and GF both perform better than GD-3D. However, no distinction is found between msKF and GF schemes. Summarizing, the overall inferences for the simulation pair H are found to be statistically inconclusive and highly sensitive, thus one cannot conclude that a specific cumulus parameterization systematically outperforms the others.

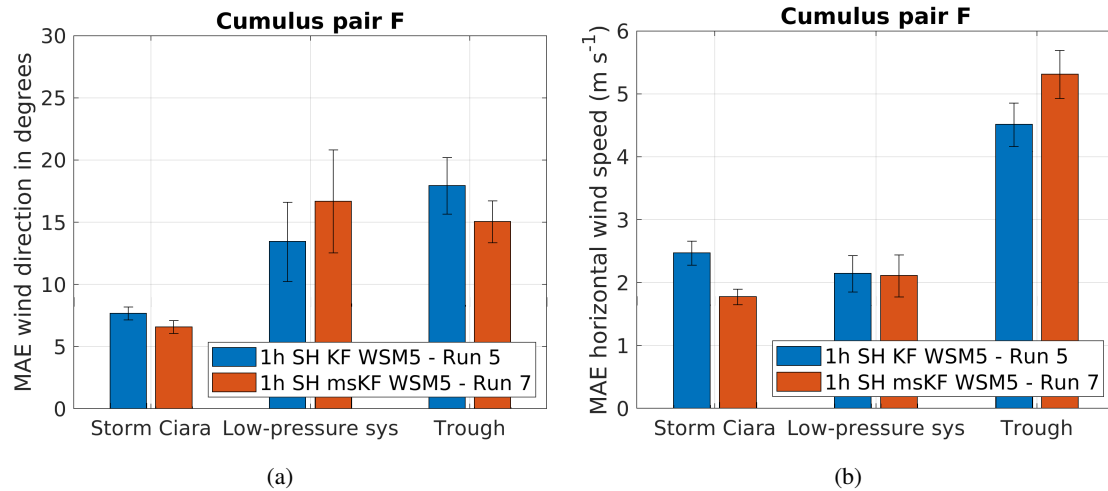


Figure 14. Performance evaluation for simulation pair F considering a change in cumulus scheme, as described in Table 2. (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.

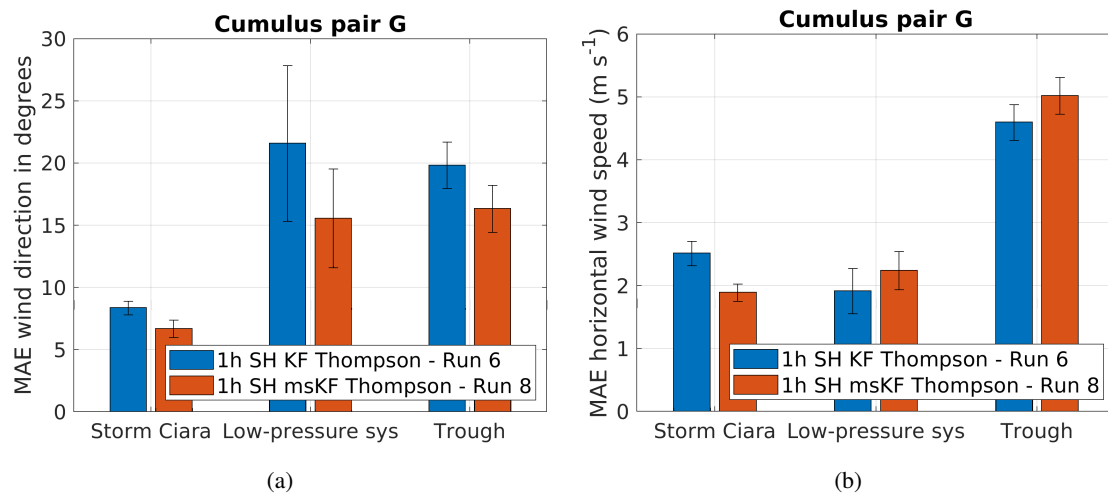


Figure 15. Performance evaluation for simulation pair G considering a change in cumulus scheme, as described in Table 2. (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.

Simulation pairs I and J focus on the impact of microphysics schemes. Starting with pair I, results are presented in Fig. 17. Pair I uses scale-aware cumulus schemes (msKF and GF) combined with high-complexity microphysics (Morrison and Thompson).

570 For precipitation, results SH PBL coupled with non-scale-aware KF cumulus schemes and varies the microphysics scheme between WSM5 and Thompson. Overall, no distinction in better performance by either microphysics schemes considering MAE of wind direction and wind speed is observed (Fig. 17a & 17b). For Storm Ciara and the trough passage EWE, wind

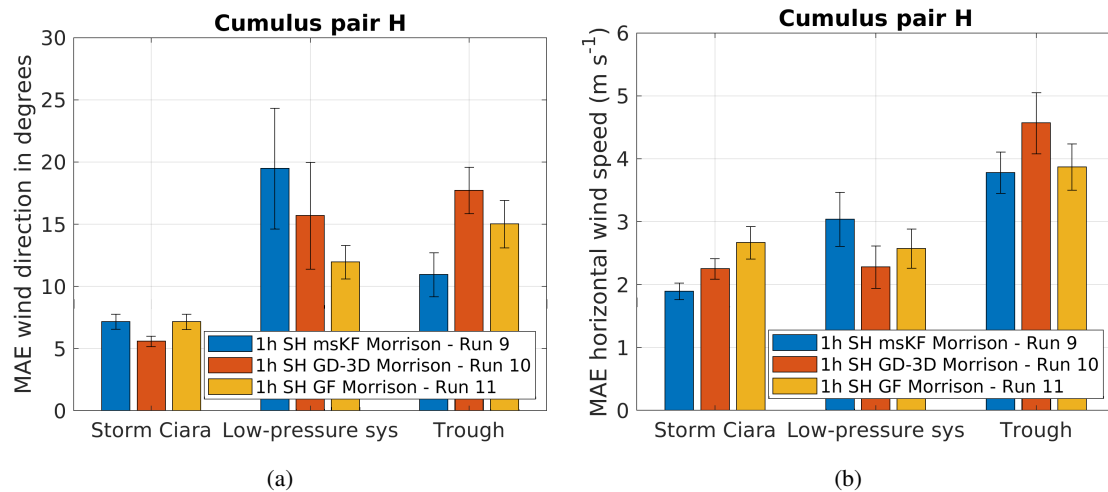
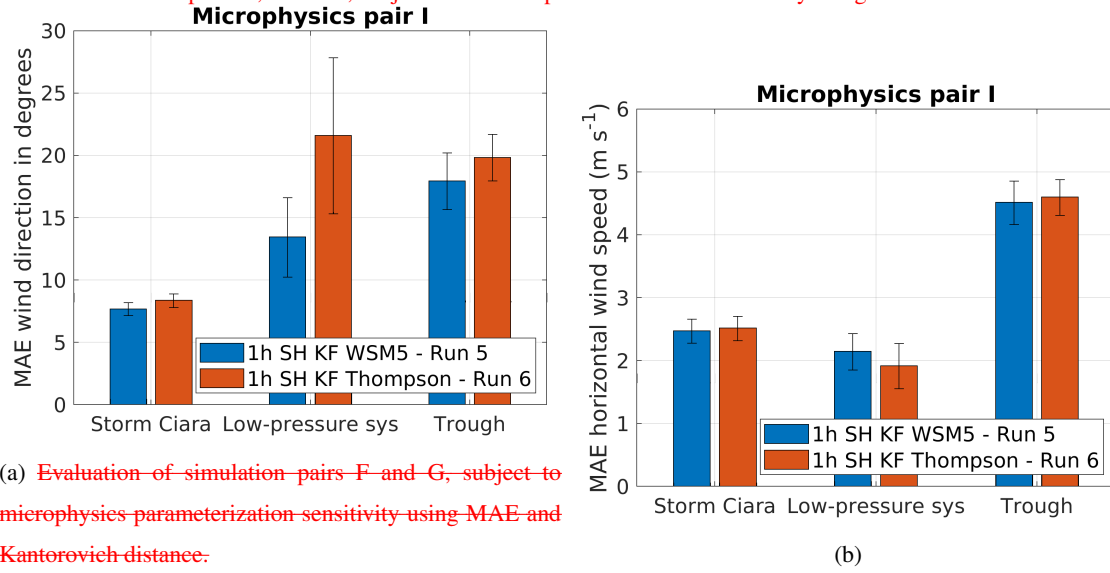


Figure 16. Performance evaluation for simulation pair G considering a change in cumulus scheme, as described in Table 2. (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.

speed and wind direction MAE are found to be highly sensitive to the combination of cumulus and microphysics setup. Varying magnitudes and time lags in the precipitation time series were produced by different combinations of cumulus and microphysics schemes, with higher magnitude observed for Morrison and Thompson microphysics coupled with scale-aware cumulus insensitive to the variation in microphysics schemes. Qualitatively, However, this trend is not clear for the case of the low-pressure system, where the Thompson scheme is found to produce different MAE of wind direction, with significantly larger error margins (Fig. 17a & 17b). However, when comparing the combination of cumulus schemes to microphysics setups, more specifically, the combination of WSM5 + KF to Thompson + KF/msKF (Fig. 14a and Fig. 15a), lower MAE of wind direction are observed for Thompson microphysics when combined with the contour plots for WRF versus RADAR observations scale-aware msKF cumulus scheme. However, this trend is reversed for WSM5 + msKF cumulus schemes. This appears to highlighting the importance of a suitable combination of cumulus and microphysics schemes.

Results for simulation pair J, which applies SH PBL with msKF cumulus schemes and varies WSM5, Thompson and Morrison microphysics schemes, are presented in Fig. ?? also result in significantly different reproductions of precipitation fronts. Once more, both quantitative and qualitative results on precipitation accuracy are inconclusive. 18. Considering MAE of wind direction (Fig. 18a), for the case of Storm Ciara and low-pressure system, a clear distinction in the performance of different microphysics schemes is not observed. For the case of trough passage, Morrison microphysics perform better in comparison. For wind speed MAE (Fig. 18b), Storm Ciara shows similar results as for wind direction. For the low-pressure system, WSM5 and Thompson produce better wind speed than Morrison. For the trough passage, Morrison outperforms WSM5 and Thompson microphysics. Overall, for simulation pair J, Storm Ciara shows an insensitivity to the variation of microphysics schemes. For the low-pressure system, no clear trend in better performance is observed, whereas a clear advantage in using the more complex Morrison scheme is observed in the trough passage case.

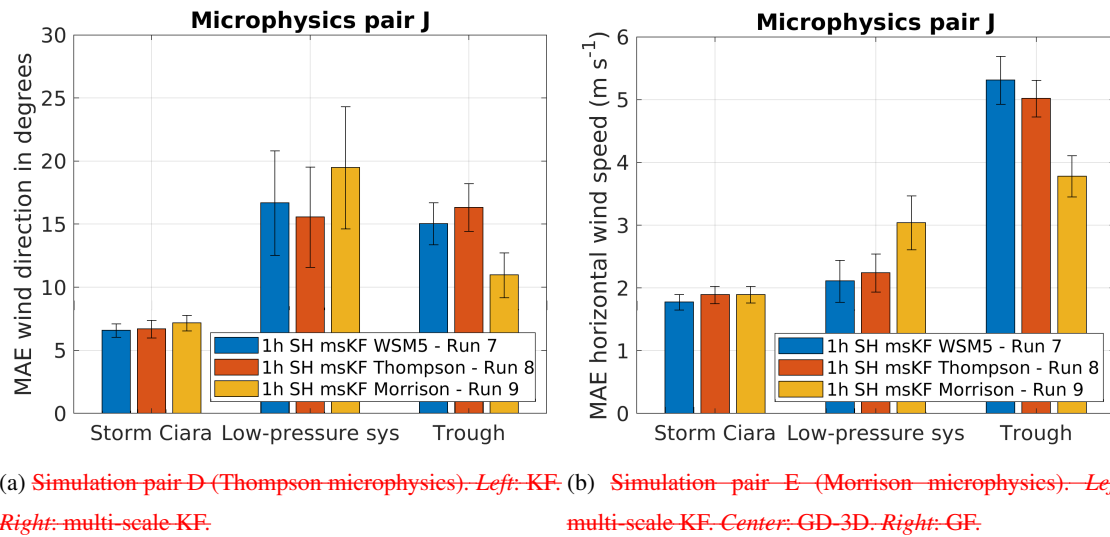
Evaluation of simulation pairs C, D and E, subject to cumulus parameterization sensitivity using MAE and Kantorovich distance.



(a) Evaluation of simulation pairs F and G, subject to microphysics parameterization sensitivity using MAE and Kantorovich distance.

(b)

Figure 17. Performance evaluation for simulation pair I considering change in microphysics schemes, as described in Table 2. (a) MAE comparison for wind direction. (b) MAE comparison for wind speed.



(a) Simulation pair D (Thompson microphysics). Left: KF. Right: multi-scale-KF. (b) Simulation pair E (Morrison microphysics). Left: multi-scale-KF. Center: GD-3D. Right: GF.

Figure 18. Contours of RADAR precipitation (contour lines in foreground) and WRF precipitation MAE comparison for wind direction (filled contours in background) (a) simulation pair D and (b) simulation pair E wind speed.

4.7 Discussion

The previous sections investigated the individual influence of varying a single physics parameterization in the modeling chain on the accuracy of the match between WRF simulation results and field data. ~~Even though a clear trend of increased fidelity observational data when subject to 3 unique case studies. A clear trend in improved performance~~ with higher model complexity ~~has been observed in Table ??, individual comparisons are more subtle~~ common to all case studies is not found.

When looking at ~~domain configurations, it has been clearly shown that larger computational domains result in significantly more short-timescale fluctuations of wind speed and direction, which closely resemble those present in the field observations. Similar results were observed in a recent study by Yu et al. (2021). Even though this produces qualitatively better time series~~ the update interval of LBC, the qualitative differences in hourly and 3-hourly update intervals are found to be marginal. The quantitative indicators show a unanimous improvement for the case of Storm Ciara. (see Fig. ??), this does not always translate into an unambiguous improvement of quantitative MAE and Kantorovich metrics considered here (see Table ??). In contrast, for 8 & 9). However, for the low-pressure system and the trough passage this distinction is not so evident.

The variation in PBL scheme results in highly sensitive metrics for all 3 case studies. For the case of Storm Ciara, a clear advantage in using scale-aware SH PBL in comparison to non-scale-aware MYNN PBL is observed (see Figs. 10 & 11). However, this trend is not evident for the case of the temporal resolution of the lateral boundary conditions, where the qualitative improvement of higher resolution is somewhat more subtle (see Fig. ??), low-pressure system and the quantitative indicators show a unanimous improvement (see Table ??). These observations advocate the use of multiple quantitative and qualitative metrics in sensitivity studies, as well as motivate the development of more advanced metrics to capture the match between simulations and field observations.

For wind speed and direction trough passage. When comparing scale-aware schemes of different fidelity, more specifically Zhang PBL and SH PBL, a promising trend is observed in which scale-aware PBL and cumulus schemes lead the higher complexity Zhang PBL leads to a better match with SCADA data for the current storm Ciara case. Based on the average NED, simulation case 12 is found to be the most performant for the current storm Ciara case study. Albeit case 12 has the lowest wind direction NED of all considered simulations, it is outperformed by several other setups when focusing on wind speed, and a different evaluation metric might have led to a different simulation ranking. That being said, in terms of wind speed and direction metrics, a clear and significant difference is found between simulations that employ both scale-aware PBL and cumulus schemes (cases 8–12) and those that do not (cases 1–7). This justifies the further development and application of scale-aware physics parameterizations. trough passage case (Fig. 13). However, this trend is not observed for Storm Ciara and the low-pressure system. Furthermore, considering wind speeds, the Zhang PBL run is either the best setup (cold front) or results in MAE very close to the best setup (Storm Ciara and low-pressure system). In contrast, the Zhang PBL setup results in higher wind direction errors for Storm Ciara and the low-pressure system cases.

Results for precipitation are much more inconclusive Regarding the cumulus and microphysics simulation pairs, the combination of cumulus and microphysics is observed to have more impact on MAE of wind direction in comparison to variation in stand-alone cumulus or microphysics schemes. This is highlighted by Storm Ciara and the low-pressure system, where the

change in microphysics schemes in combination with msKF cumulus results in marginal changes in MAE (see Fig. 18). However, when comparing the combinations of lower order microphysics with scale-aware and non-scale-aware cumulus schemes for wind direction, i.e. no clear tendency towards higher accuracy with increased model complexity is found. Furthermore, the time series, WSM5/Thompson + KF/msKF, results indicate an overall reduction in MAE for Thompson + msKF (see Fig. 14a & 15a). This observation potentially indicates scale-aware cumulus schemes to be more compatible with higher-order microphysics schemes. The performance of cumulus and microphysics schemes is found to be strongly dependent on the type of weather phenomenon.

A qualitative perspective on precipitation indicates all WRF simulations to be highly sensitive to the combination of physics schemes and type of EWE. The qualitative analysis yielded little to no conclusions on the precipitation modeling fidelity of the considered WRF physics setups in this study. As an example, the results for the case of Storm Ciara are presented in Fig. ?? and the snapshots in 19. A direct quantitative comparison of simulated reflectivity and observed raw radar fields using, e.g., Fig. ?? show RADAR reflectivity and associated precipitation produced in different runs to have a very wide spread in reproduction, indicating a strong sensitivity to model setup. Interestingly, the ensemble of all considered simulations ranks second in terms of precipitation accuracy, whereas only fifth in terms of average wind NED metric (see Table ??). These considerations promote the use of ensemble techniques including data assimilation for precipitation modeling. Tools for comparing gridded observations such as MODE (Newman et al., 2022) is impeded by the lack of filtering and post-processing information on the latter raw data. Therefore, a quantitative assessment of precipitation modeling is out of scope of the current paper and left for future work.

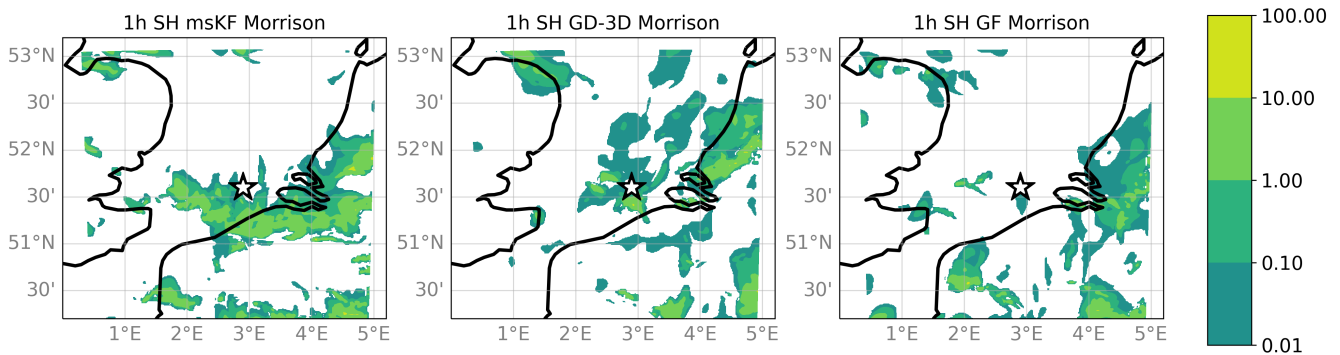


Figure 19. Contours of WRF precipitation rate in mm h^{-1} for the case of Storm Ciara on 10 February 2020 at 04:40 UTC. The plots are presented for cumulus simulation pair H for domain d04. The star in the plots represents the offshore wind farm of interest.

Finally, the ensemble average (as defined in Sect. 3.3) is observed to rank very similar to the best-case model setup (see Table 3 & 4) for the cases of Storm Ciara and the low-pressure system. However, the fast changes in wind direction are dampened by the ensemble averaging (see Fig. 5 & 6). For the trough passage, the ensemble averaging performs poorly compared to the best-case setup by a significant margin (Table 5), caused by a persistent offset by all but the best-case setup.

5 Conclusions and recommendations

The complexity in determining ~~the optimal~~ an optimal combination of physics setup for the operational use of the ~~WRF-ARW model in~~ WRF model in the frame of wind energy applications ~~has been well established by literature to exhibit strong dependencies on location-specific weather events and the combination of various physics parameterizations. In this study, a multi-variant sensitivity analysis is performed for the extreme weather case of storm Ciara as observed over the Belgian North Sea is analyzed in this study. A multi-event sensitivity analysis for WRF NWP model is performed considering three extreme weather events: Storm Ciara on 10 February 2020 over the Belgian North Sea. The event produced~~, a low-pressure system on ~~24 December 2020 and a trough passage on 27 June 2020. These events have been identified to be potentially harmful for the operation of offshore wind farms. The events resulted in~~ fast changes in wind direction ~~which can potentially lead to leading to severe yaw misalignment of the turbines, with the potential to result in~~ significant off-design turbine ~~loading and strong wind-farm power excursions. WRF results for~~ load cases and variability in power production. This sensitivity analysis utilizes ~~operational wind farm data (SCADA) for evaluating WRF simulated~~ wind direction and wind speed ~~have been evaluated against SCADA data of an operational offshore wind farm. Precipitation results were~~ results. In addition, the simulated precipitation is qualitatively compared to ~~RADAR radar~~ data from RMI-B. This sensitivity ~~analysis explores~~ study analyses the impact of ~~temporal resolution of lateral boundary conditions, domain configuration update interval of LBC and sub-grid scale modelling modeling~~ techniques used for PBL, cumulus ~~and microphysical~~, and microphysics parameterizations.

The results of this sensitivity analysis indicate WRF simulations to be highly sensitive to the type of event and the combination of physics parameterizations. Starting with the variation in update interval of LBC, overall better performance for ~~wind variables when considering~~ hourly update interval of LBC is observed for the case of Storm Ciara. However, for the low-pressure system and trough passage cases no such trend is observed. In general, WRF simulations comprising scale-aware ~~parameterization techniques. More specifically, a clear distinction was observed between simulations that employ scale-aware schemes for both PBL~~ PBL physics schemes appear to perform better in comparison to non-scale-aware physics schemes, as the best-case ~~setups for all three events feature scale-aware PBL schemes. Concerning cumulus and microphysics parameterizations, the suitable combination of cumulus and~~ cumulus ~~schemes and those that do not. In~~ microphysics is observed to be highly dependent and sensitive to the type of weather phenomenon. The combination of schemes is observed to have more impact than a stand-alone variation for either of these events.

Overall, in view of modeling ~~SCADA wind speed and direction, the most performant combination of setup and parameterizations consists of a~~ local wind direction and wind speed at the location of the farms, three independent best-case setups are identified for the three case studies. A single best WRF model setup for both wind direction and wind speed for all three case studies is not found. For the case of Storm Ciara, the best-case setup is identified to combine scale-aware ~~Shin-Hong planetary boundary layer SH~~ PBL scheme coupled with scale-aware ~~Grell-Freitas cumulus parameterization, 6-class double moment Morrison microphysics, together with hourly boundary condition updates on extended simulation domains. The representation of precipitation fronts on the other hand is shown to be highly sensitive to model setup, msKF cumulus parameterization~~ and no quantitative trends could be observed. That said, results for the ensemble average of precipitation are promising.

5-class single moment WSM5 microphysics. For the low-pressure system, the best-case setup combines, scale-aware SH PBL, non-scale-aware KF cumulus, and WSM5 microphysics schemes. For the trough passage, the best-case setup is identified to combine, scale-aware Zhang 3D PBL, scale-aware GF cumulus, and 6-class double moment Morrison microphysics schemes. The best-case setups for all cases utilize hourly reanalysis dataset as the LBC and scale-aware PBL schemes.

~~In terms of future work, expanding the current sensitivity analysis to other high-impact weather events is an important topic to assess the generalization of the current observations. Additionally, further study on the sensitivity and adequate modeling of precipitation is necessary, including data assimilation of operational RADAR data in the simulations to improve model initialization, as well as dedicated localized precipitation measurements to provide reliable field data. Finally, an~~ interesting area of further research would be to perform similar sensitivity studies at finer sub-kilometer resolutions including recent advancements such as three-dimensional scale-aware PBL schemes ~~such as in Zhang et al. (2018) and Senel et al. (2020).~~ (Zhang et al., 2018; Senel et al., 2020). Furthermore, expanding the sensitivity analysis to include events such as a dunkelflaute (Li et al., 2021) and wind ramps (Gallego-Castillo et al., 2015) will allow a broader assessment of EWE modeling relevant to wind energy. Also, a quantitative assessment of ground-level precipitation modeling with local precipitation measurements from disdrometers and tipping buckets is of general interest to assess, e.g., the risk of leading edge erosion of wind turbine blades (Law and Koutsos, 2020).

Acknowledgements. This work has received funding from the Flemish Government through the Agency for Innovation and Entrepreneurship (Vlaams Agentschap Innoveren en Ondernemen, VLAIO) through the cSBO project SeaFD in the context of the MaDurOS program for Material Durability for Off-Shore of the cluster on Strategic Initiatives for Materials in Flanders (SIM). Furthermore, the authors acknowledge VLAIO funding through the RAINBOW project in the context of SIM and the Blue Cluster. The authors thank dr. Laurent Delobbe from the Royal Meteorological Institute of Belgium for providing ~~RADAR~~ radar reflectivity data from the dual-polarization C-band ~~RADAR~~ radar, located in Jabbeke at the Belgian ~~North-Sea~~ coast.

References

- AbuGazia, M., El Damatty, A. A., Dai, K., Lu, W., and Ibrahim, A.: Numerical model for analysis of wind turbines under tornadoes, *Engineering Structures*, 223, 111–157, 2020.
- Aird, J. A., Barthelmie, R. J., Shepherd, T. J., and Pryor, S. C.: WRF-simulated low-level jets over Iowa: characterization and sensitivity studies, *Wind Energy Science*, 6, 1015–1030, 2021.
- Arakawa, A., Jung, J.-H., and Wu, C.-M.: Toward unification of the multiscale modeling of the atmosphere, *Atmospheric Chemistry and Physics*, 11, 3731–3742, 2011.
- Bakhshi, R. and Sandborn, P.: The effect of yaw error on the reliability of wind turbine blades, in: *Energy Sustainability*, vol. 50220, p. V001T14A001, American Society of Mechanical Engineers, 2016.
- Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, *Nature*, 525, 47–55, 2015.
- Carvalho, D., Rocha, A., Gómez-Gesteira, M., and Santos, C.: A sensitivity study of the WRF model in wind simulation for an area of high wind energy, *Environmental Modelling & Software*, 33, 23–34, 2012.
- Carvalho, D., Rocha, A., Gómez-Gesteira, M., and Santos, C. S.: Sensitivity of the WRF model wind simulation and wind energy production estimates to planetary boundary layer parameterizations for onshore and offshore areas in the Iberian Peninsula, *Applied Energy*, 135, 234–246, 2014.
- Chen, X., Xue, M., Zhou, B., Fang, J., Zhang, J. A., and Marks, F. D.: Effect of Scale-Aware Planetary Boundary Layer Schemes on Tropical Cyclone Intensification and Structural Changes in the Gray Zone, *Monthly Weather Review*, 149, 2079–2095, 2021.
- Chi, S.-Y., Liu, C.-J., Tan, C.-H., and Chen, Y.-H.: Study of typhoon impacts on the foundation design of offshore wind turbines in Taiwan, *Proceedings of the Institution of Civil Engineers-Forensic Engineering*, 173, 35–47, 2020.
- Choi, H.-J. and Han, J.-Y.: Effect of scale-aware nonlocal planetary boundary layer scheme on lake-effect precipitation at gray-zone resolutions, *Monthly Weather Review*, 148, 2761–2776, 2020.
- Cunden, T. M., Dhunny, A., Lollchund, M., and Rughooputh, S.: Sensitivity Analysis of WRF Model for Wind Modelling Over a Complex Topography under Extreme Weather Conditions, in: *2018 5th International Symposium on Environment-Friendly Energies and Applications (EFEA)*, pp. 1–6, IEEE, 2018.
- Damiani, R., Dana, S., Annoni, J., Fleming, P., Roadman, J., van Dam, J., and Dykes, K.: Assessment of wind turbine component loads under yaw-offset conditions, *Wind Energy Science*, 3, 173–189, 2018.
- Doubrawa, P. and Muñoz-Esparza, D.: Simulating real atmospheric boundary layers at gray-zone resolutions: How do currently available turbulence parameterizations perform?, *Atmosphere*, 11, 345, 2020.
- Dudhia, J.: A history of mesoscale model development, *Asia-Pacific Journal of Atmospheric Sciences*, 50, 121–131, 2014.
- Efstathiou, G., Zoumakis, N., Melas, D., Lolis, C., and Kassomenos, P.: Sensitivity of WRF to boundary layer parameterizations in simulating a heavy rainfall event using different microphysical schemes. Effect on large-scale processes, *Atmospheric Research*, 132–133, 125–143, <https://doi.org/https://doi.org/10.1016/j.atmosres.2013.05.004>, 2013.
- Fujita, T. T.: Manual of downburst identification for project NIMROD., *SMRP Res. Paper*, 156, 104, 1978.
- Gallego-Castillo, C., Cuerva-Tejero, A., and Lopez-Garcia, O.: A review on the recent history of wind power ramp forecasting, *Renewable and Sustainable Energy Reviews*, 52, 1148–1157, 2015.
- García-Díez, M., Fernández, J., Fita, L., and Yagüe, C.: Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe, *QJ Roy. Meteor. Soc.*, 139, 501–514, 2013.

- 740 Giannakopoulou, E.-M. and Nhili, R.: WRF model methodology for offshore wind energy applications, *Advances in Meteorology*, 2014, 2014.
- Grell, G. A. and Dévényi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, *Geophysical Research Letters*, 29, 38–1, 2002.
- Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, 745 *Atmospheric Chemistry and Physics*, 14, 5233–5250, 2014.
- Hahmann, A. N., Vincent, C. L., Peña, A., Lange, J., and Hasager, C. B.: Wind climate estimation using WRF model output: method and model sensitivities over the sea, *International Journal of Climatology*, 35, 3422–3439, 2015.
- Hannesdóttir, Á. and Kelly, M.: Detection and characterization of extreme wind speed ramps, *Wind Energy Science*, 4, 385–396, 2019.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: 750 The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049, 2020.
- Hong, S.-Y. and Dudhia, J.: Next-generation numerical weather prediction: Bridging parameterization, explicit clouds, and large eddies, *Bulletin of the American Meteorological Society*, 93, ES6–ES9, 2012.
- Hong, S.-Y. and Lim, J.-O. J.: The WRF single-moment 6-class microphysics scheme (WSM6), *Asia-Pacific Journal of Atmospheric Sciences*, 42, 129–151, 2006.
- 755 Hong, S.-Y., Dudhia, J., and Chen, S.-H.: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation, *Monthly weather review*, 132, 103–120, 2004.
- Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, *Monthly weather review*, 134, 2318–2341, 2006.
- Huang, H., Winter, J. M., Osterberg, E. C., Hanrahan, J., Bruyère, C. L., Clemins, P., and Beckage, B.: Simulating precipitation and temper- 760 ature in the Lake Champlain basin using a regional climate model: limitations and uncertainties, *Climate Dynamics*, 54, 69–84, 2020.
- Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *Journal of Geophysical Research: Atmospheres*, 113, 2008.
- Islam, T., Srivastava, P. K., Rico-Ramirez, M. A., Dai, Q., Gupta, M., and Singh, S. K.: Tracking a tropical cyclone through WRF–ARW simulation and sensitivity of model physics, *Natural Hazards*, 76, 1473–1495, 2015.
- 765 Jeworrek, J., West, G., and Stull, R.: Evaluation of cumulus and microphysics parameterizations in WRF across the convective gray zone, *Weather and Forecasting*, 34, 1097–1115, 2019.
- Kain, J. S.: The Kain–Fritsch convective parameterization: an update, *Journal of applied meteorology*, 43, 170–181, 2004.
- Kala, J., Andrys, J., Lyons, T. J., Foster, I. J., and Evans, B. J.: Sensitivity of WRF to driving data and physics options on a seasonal time-scale for the southwest of Western Australia, *Climate Dynamics*, 44, 633–659, 2015.
- 770 Kalverla, P. C., Steeneveld, G.-J., Ronda, R. J., and Holtslag, A. A.: An observational climatology of anomalous wind events at offshore meteomast IJmuiden (North Sea), *Journal of Wind Engineering and Industrial Aerodynamics*, 165, 86–99, 2017.
- Kantorovitch, L.: On the translocation of masses, *Management science*, 5, 1–4, 1958.
- Law, H. and Koutsos, V.: Leading edge erosion of wind turbines: Effect of solid airborne particles and rain on operational wind farms, *Wind Energy*, 23, 1955–1965, <https://doi.org/https://doi.org/10.1002/we.2540>, 2020.
- 775 Li, B., Basu, S., Watson, S. J., and Russchenberg, H. W.: A Brief Climatology of Dunkelflaute Events over and Surrounding the North and Baltic Sea Areas, *Energies*, 14, 6508, 2021.
- Marshall, J. and Palmer, W.: Relation of raindrop size to intensity, *Journal of Meteorology*, 5, 165–166, 1948.

- Mooney, P., Mulligan, F., and Fealy, R.: Evaluation of the sensitivity of the weather research and forecasting model to parameterization schemes for regional climates of Europe over the period 1990–95, *Journal of Climate*, 26, 1002–1017, 2013.
- 780 Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one-and two-moment schemes, *Monthly weather review*, 137, 991–1007, 2009.
- Nakanishi, M. and Niino, H.: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog, *Boundary-Layer Meteorology*, 119, 397–407, 2006.
- Negro, V., López-Gutiérrez, J.-S., Esteban, M. D., and Matutano, C.: Uncertainties in the design of support structures and foundations for
785 offshore wind turbines, *Renewable energy*, 63, 125–132, 2014.
- Newman, K., J. Opatz, T., Jensen, J., Prestopnik, H., Soh, L., Goodrich, B., Brown, R. B., and Gotway, J. H.: MET-MODE, in: *The MET Version 10.1.0 User’s Guide*, DTC, 2022.
- Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Coen, J. L., Gochis, D. J., Ahmadov, R., Peckham, S. E., et al.: The weather research and forecasting model: Overview, system efforts, and future directions, *Bulletin of the American
790 Meteorological Society*, 98, 1717–1737, 2017.
- Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., Lara-Fanego, V., and Tovar-Pescador, J.: Analysis of WRF model wind estimate sensitivity to physics parameterization choice and terrain representation in Andalusia (Southern Spain), *Journal of Applied Meteorology and Climatology*, 52, 1592–1609, 2013.
- Senel, C. B., Temel, O., Muñoz-Esparza, D., Parente, A., and van Beeck, J.: Gray zone partitioning functions and parameterization of
795 turbulence fluxes in the convective atmospheric boundary layer, *Journal of Geophysical Research: Atmospheres*, 125, e2020JD033 581, 2020.
- Shin, H. H. and Hong, S.-Y.: Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions, *Monthly Weather Review*, 143, 250–271, 2015.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., et al.:
800 A description of the advanced research WRF model version 4, *National Center for Atmospheric Research: Boulder, CO, USA*, 145, 145, 2019.
- Solari, G.: Thunderstorm Downbursts and Wind Loading of Structures: Progress and Prospect. *Front, Built Environ*, 6, 63, 2020.
- Stergiou, I., Tagaris, E., and Sotiropoulou, R.-E. P.: Sensitivity assessment of WRF parameterizations over Europe, in: *Multidisciplinary digital publishing institute proceedings*, vol. 1, p. 119, 2017.
- 805 Tewari, Mukul, N., Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., et al.: Implementation and verification of the unified NOAA land surface model in the WRF model (Formerly Paper Number 17.5), in: *20th conference on weather analysis and forecasting/16th conference on numerical weather prediction*, pp. 11–15, 2004.
- Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization, *Monthly Weather Review*, 136, 5095–5115, 2008.
- 810 Wang, Y. and Basu, S.: Utilizing the Kantorovich metric for the validation of optical turbulence predictions, *Optics letters*, 41, 4008–4011, 2016.
- Wilks, D. S.: *Statistical methods in the atmospheric sciences*, vol. 4, Elsevier, 2019.
- Wyngaard, J. C.: Toward numerical modeling in the “Terra Incognita”, *Journal of Atmospheric Sciences*, 61, 1816–1826, 2004.
- Xu, H., Wang, Y., and Wang, M.: The performance of a scale-aware nonlocal PBL scheme for the subkilometer simulation of a deep CBL
815 over the Taklimakan Desert, *Adv. Meteor.*, 8759594, 2018.

- Yu, Z., Wu, M., Min, J., Yan, Y., and Lou, X.: Impacts of WRF Model Domain Size on Meiyu Rainfall Forecasts over Zhejiang, China, *Asia-Pacific Journal of Atmospheric Sciences*, pp. 1–16, 2021.
- Zhang, X., Bao, J.-W., Chen, B., and Grell, E. D.: A three-dimensional scale-adaptive turbulent kinetic energy scheme in the WRF-ARW model, *Monthly Weather Review*, 146, 2023–2045, 2018.
- 820 Zheng, Y., Alapaty, K., Herwehe, J. A., Del Genio, A. D., and Niyogi, D.: Improving high-resolution weather forecasts using the Weather Research and Forecasting (WRF) Model with an updated Kain–Fritsch scheme, *Monthly Weather Review*, 144, 833–860, 2016.