



# Article Accuracy Assessment of WRF Model in the Context of Air Quality Modeling in Complex Terrain

Mateusz Rzeszutek <sup>1,\*</sup>, Adriana Kłosowska <sup>2</sup> and Robert Oleniacz <sup>1,\*</sup>

- <sup>1</sup> Department of Environmental Management and Protection, Faculty of Geo-Data Science, Geodesy and Environmental Engineering, AGH University of Science and Technology, Mickiewicza 30 Av., 30-059 Krakow, Poland
- <sup>2</sup> Municipality of Krakow, Climate-Energy-Water Management, Os. Szkolne 27, 31-977 Krakow, Poland; adriana.szulecka@gmail.com
- \* Correspondence: rzeszut@agh.edu.pl (M.R.); oleniacz@agh.edu.pl (R.O.)

Abstract: Output data from the Weather Research and Forecasting (WRF) model are frequently used in air quality modeling for scientific, practical and regulatory purposes. Therefore, it is crucial to determine whether the accuracy of WRF predictions is suitable for application in air quality models on a local scale (<50 km) and in complex terrain. The presented research is unique because, to assess the accuracy of the WRF model, data from experimental data sets for the assessment of air quality models were used, which contained information about the actual conditions of selected meteorological parameters along the vertical profile of the atmosphere. The aim of the study was to conduct an evaluation of the WRF model using data derived from three field experiments designated to conduct air quality model evaluation studies for models such as AERMOD, ADMS or CALPUFF. Accuracy evaluation was carried out in relation to the grid resolution, station location (on-site and weather airport) and vertical profile of the atmosphere. Obtained results of the evaluation for temperature, wind speed and direction were analyzed with regard to the possibilities of application in air quality modeling systems. It was stated that the use of a grid with a resolution of 1 km generally resulted in statistically significantly lower values of errors for wind speed compared to a 4 km resolution. The outcomes of simulations for temperature and wind speed were sensitive with regard to the location. In on-site locations (complex terrain) significantly higher values of prediction errors (MB, MGE, RMSE) were obtained compared to the standard weather station locations (airport). In addition, wind speed predictions in on-site locations were generally biased (overestimated). Along the vertical profile of the atmosphere, up to the altitude of 100 m a.g.l., statistically significantly different outcomes of accuracy evaluation were achieved for wind speed and direction. Considering the above, caution should be exercised when using data from meteorological simulations in air quality modeling.

**Keywords:** meteorological model; WRF; accuracy assessment; grid resolution; complex terrain; air pollution modeling

## 1. Introduction

Accurate air quality predictions on a local scale (<50 km) and in complex terrain constitute one of the scientific challenges due to, among others, the necessity of replicating local meteorological conditions while maintaining adequate computational efficiency. The fundamental problem associated with air quality modeling for impact assessments is the availability of representative meteorological data for the considered area, as noted by Tartakovsky et al. [1], Abril et al. [2], Karthick and Devadoss [3] or Ottosen et al. [4]. In many cases, the application of data from existing monitoring systems of meteorological parameters is not adequate. Spatial representativeness of the measurement data is limited. The use of meteorological data from sites located several to dozens of kilometers from the source may lead to unrepresentative results of dispersion modeling, especially in



Citation: Rzeszutek, M.; Kłosowska, A.; Oleniacz, R. Accuracy Assessment of WRF Model in the Context of Air Quality Modeling in Complex Terrain. *Sustainability* **2023**, *15*, 12576. https://doi.org/10.3390/ su151612576

Academic Editors: Md Firoz Khan, Ram Lal Verma and Nadim Khandaker

Received: 27 June 2023 Revised: 7 August 2023 Accepted: 17 August 2023 Published: 18 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). complex terrain [5,6]. Found in the literature, one of the solutions to this problem is to obtain data from an advanced meteorological model, e.g., WRF (Weather Research and Forecasting model) or MM5 (Fifth-Generation Penn State/NCAR Mescoscale Model). The importance of meteorological modeling in regional air quality assessment was emphasized by, among others, Sistla et al. [7], Gilliam et al. [8], Pielke and Uliasz [9], and McNider and Pour-Biazar [10].

Two of the most recognizable mesoscale meteorological models used for providing input to the air pollution models are the WRF [11] and MM5 [12]. However, the MM5 model is no longer supported and developed since 2005 and the results of research included in [1] provide evidence that for the purpose of air quality modeling the application of the simulation outputs from WRF is more effective. WRF was designed for general atmospheric research purposes as well as for operational forecasting. Its applications include, among many others, short-term meteorological forecasting [13,14], climatologic research [15,16], prediction of extreme events such as hurricanes [17,18], wind power forecasting [19,20] and air quality assessment and forecasting [2,5,21-30]. Apart from regional applications, the results from the WRF model are implemented in air quality modeling systems on a local scale (<50 km) for air quality impact assessment, e.g., [3,5,20–24] in the areas of urban agglomerations in order to assess air quality, determine the location of air quality monitoring stations and to quantify the contribution of each emission source in air quality deterioration [31–37]. Many of the aforementioned studies focused on conducting an air quality model assessment, but meteorological model assessment was occasionally neglected. In effect, conclusions from the studies are limited as they do not provide information about the accuracy of the input meteorological data.

According to the United States Environmental Protection Agency (U.S. EPA) Revisions to the Guideline on Air Quality Models [38], wind direction measurement error of 5–10° results in uncertainty ranging from 20% to 70% of the concentration determination (in a given time and space) by Gaussian plume models. It means that relatively small modification of the meteorological parameter leads to exponential increase in the error of the surface air pollutant concentration. This effect is always crucial, since the application of meteorological data from weather prediction modeling systems in the process of air quality modeling may yield unrepresentative outcomes and lead to inadequate conclusions. It can also compensate for the errors in air quality predictions, which in turn prevent the actual assessment of both the structural and parametric uncertainty of the model. This issue is particularly significant on a local and micro scale, where reliable representation of meteorological conditions is essential and there are no on-site measurements available. Publicly accessible ISD [39] and IGRA [40] meteorological parameter measurements are generally conducted at the airports located on the cities' outskirts [4]. This problem was also addressed in [41], where default wind speed correction factor (f\_roof) in the OSPM model was responsible for the underestimation of PM<sub>10</sub> and PM<sub>2.5</sub> particulate matter concentration predictions. The use of meteorological data from weather prediction models in air quality modeling (including radioactive pollutants) is therefore unavoidable on a local scale, which is confirmed by numerous studies that used data from meteorological models [21,23,35,42–47]. On the other hand, many studies [2,5,6,48] demonstrated that the use of data from meteorological models results in decreased accuracy in predicted air pollutant concentration levels.

The above is reflected in the U.S. EPA Revisions to the Guideline on Air Quality Models [38]. The use of meteorological data from weather prediction models is allowed, but only when representative measurement data are not available or conducting meteorological measurements is not feasible. The application of these data must be preceded by an accuracy assessment. Noting the research on the evaluation of dispersion models on a local scale, including complex terrain, it should be emphasized that regardless of the study area, the best air quality predictions were obtained when a simultaneous measurement campaign of the meteorological parameters was conducted in the vertical profile of the atmosphere [5,6,49]. This implies that the extrapolation methods along the vertical profile

of the atmosphere included in the air quality models for selected meteorological parameters do not yield reliable results. Therefore, it is particularly significant to pay attention to the accuracy of the representation of the vertical profile of meteorological parameters by weather prediction models, since the acquisition of such data for impact assessments of emission sources on air quality is expensive and probably not economically justified in many situations.

The aim of the study was to evaluate the accuracy of the WRF model in the scope of basic meteorological parameters used in the air quality modeling process (temperature, wind speed and direction) at different levels above the ground level on a local scale and complex terrain. The paper evaluated whether an increase in the density of the computational grid from 4 km to 1 km resulted in more accurate outcomes of the meteorological parameter predictions as, according to [50,51], an increase in the grid resolution does not always yield better wind field predictions. The accuracy assessment results of selected meteorological parameters were compared with regard to different locations of the sites (synoptic, on-site). It allowed for the examination of the stability of the weather prediction model subjected to varying local terrain conditions. The research experiment was designed in a way that allowed for the application of conducted simulations for other purposes, such as evaluation of air quality modeling systems using meteorological data from WRF simulations or parameterization of air quality models with regard to input meteorological data. Therefore, it was decided to apply three experimental dispersion model assessment datasets made available by the U.S. EPA [52]. These experiments are standardly used to evaluate the accuracy of the development versions of AERMOD and ADMS models. The choice of multiple experiments was also motivated by the conclusions in [53], where it was shown that it is not always feasible to obtain homogenous conclusions for the same air quality model with an analogous configuration for different study areas that have common features relevant to air quality modeling.

#### 2. Materials and Methods

#### 2.1. Study Area and Data Completeness

Study areas consisted of three experimental datasets included in the Model Evaluation Databases (MED), i.e., Martin Creek (MC), Lovett (LV) and Tracy Power Plant (TR). These databases were originally made available by the U.S. EPA [52] for the evaluation and development of air pollutant dispersion models used for regulatory purposes. Results of studies focused on the evaluation of dispersion models such as CALPUFF and AERMOD with the use of the aforementioned datasets were presented, among others, in [5,53–57]. These areas of study were selected because they are characterized by complex terrain. As part of the experiments, measurements of selected meteorological parameters were conducted in the vertical profile of the atmosphere. Graphical representation of the study areas is shown in Figure 1, and the characteristics of the meteorological sites are presented in Table 1.

Table 1. Basic information concerning meteorology	ogical stations providing measure	ments for the WRF
model accuracy evaluation.		

MED	Station *	UASF-WBAN	Latitude [°]	Longitude [°]	Elevation [m a.s.l.]
МС	ONS LVIA	n/d 725,170-14,737	40.790 40.651	$-75.139 \\ -75.449$	89.0 118.9
LV	ONS	n/d	41.240	-73.982	69.0
	WCA	725,037-94,745	41.067	-73.708	115.5
	SIA	725,038-14,714	41.500	-74.100	149.7
TR	ONS	n/d	39.561	-119.51	1340.3
	RTIA	724,880-23,185	39.484	-119.771	1344.2

\* ON—Son-site (site—specific location meteorological station); LVIA—Lehigh Valley International Airport, WCA—Westchester County Airport; SIA—Stewart International Airport; RTIA—Reno/Tahoe International Airport.



**Figure 1.** Graphical representation of the study areas (MC—Martin Creek, LV—Lovett, TR—Tracy Power Plant) with the borders of computational domains set in the WRF model with the resolutions of 4 km (WRF4) and 1 km (WRF1).

As part of each experiment, measurements of temperature, wind direction and speed at various altitudes were carried out (Figure 2). The Martin Creek experiment was conducted from 1 May 1992 to 19 May 1993 on the Pennsylvania/New Jersey border. Measurements of temperature, wind speed, wind direction at the altitude of 10 m a.g.l. were recorded from an instrumented tower, and the hourly measurements of wind speed at multiple heights were conducted using sodar [54,57,58]. The Lovett experiment was conducted from December 1987 to December 1988 in a rural area of New York State. Measurements of meteorological parameters were conducted at the heights of 50, 100 and 150 m a.g.l. from the meteorological tower [58,59]. The Tracy Power Plant experiment was carried out in Nevada State, approximately 27 km from Reno. It was part of the EPA Complex Terrain Model Development (CTMD) program [60]. Measurement period covered 128 h in August 1984, consisting of 14 measurement sub-periods during which stable atmospheric conditions dominated. Wind speed and direction measurements were conducted to the altitude of 150 m a.g.l. from an instrumented tower. Wind measurements were extended above 150 m using Doppler acoustic sounder. Temperature measurements were extended with a tethersonde [58,61]. In addition to the on-site meteorological data (site-specific location meteorological station for experiments) there were also weather stations within computational domains, from which measurements are available through the ISD database [39]. Measurement meteorological data from the MED (ONS for each experiment) and ISD (LVIA, WCA, SIA, RTIA) databases were characterized by a time resolution of 1 h.

Results of the analysis verifying completeness of meteorological data derived from standard meteorological stations (ISD) and experiment-specific measurements in the vicinity of the emission source (ONS) are presented in Table 2 and Figure 2, respectively. Data from the LVIA (MC) and RTIA (TR) stations were characterized by a high completeness at the level of minimum 90%. In the case of the Lovett experiment, completeness of data from weather stations was visibly lower, especially for wind direction records manifested by the availability of data at the level of 59% and 83% for the SIA and WCA stations, respectively. For the on-site measurements, conducted as part of field experiments, the completeness of data was significantly higher. In the case of measurements carried out at the altitude of 10 m a.g.l. data availability ranged from 98% to 100%. However, the completeness of data decreased with the increase in altitude to 40-50% at the highest measurement levels, depending on the experiment and meteorological parameter. It should be noted that, accordingly to Figure 2, in the case of the MC experiment there were no temperature measurements in the vertical profile of the atmosphere available. In the TR experiment wind speed was not measured at the heights of 15 m a.g.l. and 50 m a.g.l. and the wind direction was not available for the level of 75 m a.g.l.

Analyzed experiments are not homogenous. First of all, data from multiple levels along the vertical profile of the atmosphere are available for the TR experiment, but the measurement collection period was exceptionally short with 128 h of total data availability. For other experiments the measurement periods were at least one year long, but for the LV the data was collected only from the levels of 10, 50 and 150 m a.g.l. All experiments concern areas characterized by complex terrain, but they are strongly differentiated in terms of the measurement campaign period or the number of data collected. Therefore, caution should be exercised when comparing the results of accuracy assessment of the WRF model between experiments.

MED Station	Completeness of Observations [%]			Number of Missing Observations			
	Station	Wind Speed	Wind Direction	Temperature	Wind Speed	Wind Direction	Temperature
MC	LVIA	100%	93%	100%	0	617	0
LV	WCA	94%	83%	95%	501	1531	483
LV	SIA	84%	59%	83%	1480	3569	1439
TR	RTIA	100%	96%	100%	0	24	0

**Table 2.** Completeness of observations for wind speed, wind direction and temperature for selected meteorological stations derived from the Integrated Surface Database (ISD).

## 2.2. Model Setup and Data Processing

The WRF v3.8 model runs were initialized using the North American Regional Reanalysis (NARR) dataset (ds608.0) developed by the National Centers for Environmental Prediction (NCEP). The NARR model assimilates (3DVAR) observational data such as temperatures, winds and moisture from radiosondes or pressure data from surface observations as well as other measurements, including precipitation [62]. The output is available with a 3-h timestamp at 29 Eta pressure levels for the period of 1979–2023 and continuing. The land use information for the WRF runs was derived from the National Land Cover Data 1992 (NLCD 1992) with a native resolution of 30 m, which was assumed representative as all the experiments were conducted during the period of 1984–1993. Recent studies show that the impact of land use on the WRF model performance is not negligible but subtle [63,64]. However, urban areas tend to block incoming solar radiation and wind profiles which may impact the temperature, wind speed and direction [65] and their adequate representation is crucial for the model's performance in highly urbanized regions [66]. Therefore, the land use data from 1992 applied in the study should not be the main determinant of the differences between the obtained simulation results and observations.

Depending on the experiment, three (MC, LV) or four (TR) computational domains were set. The coarsest domain in the TR experiment had a resolution of 36 km and in both MC and LV 24 km. In all cases the innermost, one-way nested domains had the resolutions of 4 km and 1 km, as presented in Figure 1. Nesting was carried out using 3–6 grid ratio, since [67] found no significant differences between ratios in the range of 3–7 in relation to the model performance. Simulations were carried out at 32–43 pressure levels depending on the experiment. In order to better represent the boundary layer in regions with complex terrain, a higher number of levels of pressure was used compared to the studies presented in [68]. Details concerning domain setup in WRF model for different experiments are summed up in Table 3. Simulations were carried out for time periods corresponding to the specific experiments described in Section 2.1 with a 12-h model spin-up to develop smaller-scale convective features as noted by [69].



**Figure 2.** Graphical representation of percentage share of missing observations for selected meteorological parameters measured at the ONS (on-site) stations for MC, LV and TR experiments (the numerical values in the headers indicate the height in meters above ground level).

The WRF physics configuration for the experiments was set based on the literature findings and summarized in Table 4. Boundary layer physics were resolved using the

YSU scheme [70], which was found to perform relatively better in terms of reproducing vertical wind speed profile and the 10-m wind speed compared to other schemes [71]. The microphysics, on the other hand, was resolved using the WSM6 scheme [72]. For the cumulus parameterization the Kain-Fritsch (new Eta) scheme [73] was selected, with no cumulus physics in the inner domains. The Noah LSM [74] selection for the land surface model was based on its ability to predict near-surface water mixing ratios [75], however it tends to miscalculate the boundary layer evolution in the morning hours [76]. For the longwave and shortwave radiation, the Rapid Radiative Transfer Model [77] and Dudhia [78] schemes were selected, respectively. The MM5 Dudhia scheme was found to perform slightly better in the terms of wind direction prediction than the Goddard and GFDL schemes [75].

Experiment	Domain	Size	Resolution [km]
	d01	74 imes74	24
MC	d02	72  imes 72	4
	d03	72  imes 72	1
	d01	74 imes74	24
LV	d02	72  imes 72	4
	d03	72  imes 72	1
	d01	59  imes 59	36
TD	d02	57 imes48	12
IK	d03	60 imes 48	4
	d04	$52 \times 52$	1

Table 3. Details concerning domains settings in WRF simulations for different field experiments.

Table 4. Configuration options for the WRF v3.8 modeling system.

Parameter	Selected Scheme/Data				
PBL scheme—Surface Layer scheme	Yonsei State University scheme				
Microphysics	WSM6 scheme				
Land-surface model	Noah LSM				
Sea Surface Temperature (SST)	Time varying				
Longwave Radiation Scheme	RRTM scheme				
Shortwave Radiation Scheme	Dudhia scheme				
Meteorological input fields	NCEP North American Regional Reanalysi (NARR)				
Land use data	NLCD 1992				

Extraction of the WRF model simulation results was carried out using MMIF [79], which is designed to process meteorological data from WRF and MM5 for use in air quality models such as AERMOD, SCICHEM and CALPUFF. Application of this tool was necessary for the acquisition of meteorological data at defined altitudes a.g.l. WRF model simulates values of meteorological parameters on given pressure levels; therefore, for the comparison of outcomes in the vertical profile of the atmosphere it was essential to determine simulation values at given heights relative to the ground level. According to Jiménez and Dudhia [68,80], careful selection of grid points to assess model accuracy is important. In this case, the closest grid points were selected because the variation in terrain elevation within about 1 km of the station location was small. Terrain mapping for grid resolution ranging from 100 to 4000 m is presented in [53].

#### 2.3. Methodology for Evaluating the Accuracy of WRF Model

The evaluation of the WRF model simulation results was conducted for temperature, wind speed and direction. For this purpose, a set of commonly used environmental model performance measures were applied, i.e., Mean Bias (MB), Mean Gross Error (MGE), Normalized Mean Bias (NMB), Normalized Mean Gross Error (NMGE), Root Mean Square

Error (RMSE), Index Of Agreement (IOA). Equations for calculation of the model performance measures are presented below [81–83].

$$MB = \frac{1}{n} \sum_{i=1}^{n} (M_i - O_i),$$
(1)

$$MGE = \frac{1}{n} \sum_{i=1}^{n} |M_i - O_i|,$$
 (2)

RMSE = 
$$\sqrt{\frac{\sum_{i=1}^{n} (M_i - O_i)^2}{n}}$$
, (3)

NMB = 
$$\frac{\sum_{i=1}^{n} (M_i - O_i)}{\sum_{i=1}^{n} O_i}$$
, (4)

NMGE = 
$$\frac{\sum_{i=1}^{n} |M_i - O_i|}{\sum_{i=1}^{n} O_i}$$
, (5)

$$IOA = \begin{cases} 1 - \frac{\sum_{i=1}^{n} |M_{i} - O_{i}|}{c\sum_{i=1}^{n} |O_{i} - \overline{O}|}, & when \quad \sum_{i=1}^{n} |M_{i} - O_{i}| \le c\sum_{i=1}^{n} |O_{i} - \overline{O}| \\ \frac{c\sum_{i=1}^{n} |O_{i} - \overline{O}|}{\sum_{i=1}^{n} |M_{i} - O_{i}|} - 1, & when \quad \sum_{i=1}^{n} |M_{i} - O_{i}| > c\sum_{i=1}^{n} |O_{i} - \overline{O}| \end{cases}$$
(6)

Model performance measures such as MB, MGE, RMSE return values in units of the evaluated parameter. A value of 0 indicates an ideal model. In the case of MB, caution should be exercised, as the value of 0 might mean that the negative and positive bias of predictions towards observations have balanced out. In contrast, for the NMB and NMGE a value of 0 always indicates an ideal model. Normalized values of systematic errors were applied in comparative analysis of selected parameters involving assigned groups where mean observed values differed, e.g., the change in the mean wind speed value with the altitude. In this case, Index of Agreement (IOA) indicates refined index of model performance, which was introduced in [83]. IOA is a dimensionless parameter with value in the range from -1 to 1, where 1 indicates an ideal model. The value of 0 means that the sum of model errors equals the sum of observations deviations in relation to their mean value. The value of 0.5 denotes that the sum of model errors constitutes half of the observed values deviations in relation to the mean value [82].

Emery et al. [81] proposed sets of benchmarks for the abovementioned statistical measures of the model performance for various atmospheric parameters from meteorological models used in air quality modeling. Benchmarks were proposed for wind speed (RMSE  $\leq 2 \text{ m/s}$ ; MB  $\leq \pm 0.5 \text{ m/s}$ ; IOA  $\geq 0.6$ ), wind direction (MGE  $\leq 30^{\circ}$ ; MB  $\leq \pm 10^{\circ}$ ) and temperature (MGE  $\leq$  2 K; MB  $\leq$   $\pm$ 0.5 K; IOA  $\geq$  0.8). Above values were determined based on multiple different PSU/NCAR simulations with MM5 mesoscale model but have been applied for WRF model as well [84–86]. In accordance with the interpretation included in [86] the above reference values are to be considered as guidelines rather than final tests of acceptance or rejection of the model. Outcomes of the meteorological evaluation may depend on time and location, which is de facto the subject of this study. Moreover, determination of the model evaluation criteria based on non-normalized parameters, only in the units of evaluated parameters, creates certain discomfort in terms of scientific correctness. It is difficult to assume that the accuracy of the model at the level of 1 m/s is equally relevant where the mean wind speed values and the sample variances at diverse locations are statistically significantly different. Application of homogenous criteria is not always justified, but they should be taken into account and the analysis results should be interpreted cautiously. The abovementioned guidelines were developed for the MM5 model; however they are used as a reference point for assessing the performance of meteorological models [76,87]. What is more, in accordance with U.S. EPA [88] these benchmarks should not be applied as

indicators of model simulation acceptance. The results of statistical evaluation might point to the occurrence of specific, frequently local effects that require further examination. In this study, the evaluation of meteorological predictions was divided into 3 stages. In the first stage, outcomes from computational domains of different horizontal resolutions were compared. Deciding in the first place which resolution of the computational grid produces significantly lower deviations of meteorological parameters in relation to the observations assisted in narrowing down the range of presented outcomes in the next stages.

Second stage focused on the analysis of the results obtained for meteorological measurements carried out at the ground level. In this case, measurement data from field experiments (ONS), as well as synoptic stations (ISD) were available. The ONS sites, compared to the synoptic stations (ISD), were generally located in river valleys surrounded by hills. Therefore, the ONS sites were located in the areas defined as complex terrain according to the definition in [38]. In those locations, terrain effects influence greatly the characteristics of wind field, as well as transport and diffusion of air pollutants. Based on the observations concerning data completeness (see Section 2.1) in this stage the number of observations was limited in order to maintain an equal amount of observations between compared stations. In the case of Lovett experiment, two ISD stations (WCA and SIA) were located within the computational domain; therefore, for each station data comparison was conducted separately, with two distinct datasets characterized by different numbers of observations.

Third stage involved the evaluation of the WRF model along the vertical profile of the atmosphere and was limited to the on-site stations. Analysis focused on the comparison of accuracy of the meteorological predictions with increasing altitude. This stage is of significant importance due to the possibility of implementing the data in air quality modeling. In typical air quality models applied for environmental impact assessment it is generally necessary to determine wind speed at the height of flue gases discharge point, increased by the plume rise. Values of meteorological parameters in the atmospheric vertical profile are most commonly estimated based on meteorological data at ground level, measured at standard ISD stations and atmospheric soundings, which are characterized by the measurement frequency of 2–6 times per day.

For the purpose of examining whether there were differences between analyzed cases at individual stages, the model comparison measure (MCM) was applied, defined as:

$$MCM = MGE_{model \ A} - MGE_{model \ B} \tag{7}$$

In order to verify whether the differences were statistically significant, confidence intervals (CI) at the confidence level of 95% were determined. CI intervals were estimated using non-parametric method of bootstrap available through the boot package [89,90] for the R programming language [91] in the RStudio integrated environment [92]. The methodology proposed above for comparison of the model performance outcomes was derived from the air quality model evaluation procedures described in [93,94]. The Supplementary Materials section of this paper contains a supplement to the accuracy assessment of the WRF model as well as a graphical presentation of the results. Graphical methods and selected algorithms were taken from the presentation of research results [81,95].

## 3. Results and Discussion

## 3.1. Grid Resolution

The results of the WRF model accuracy evaluation concerning two set computational grid resolutions with regard to the assessed meteorological parameters, experiments and sites are shown in Table 5. MGE and RMSE measures for wind speeds are similar regardless of the experiment and site location for the grid resolutions of 1 km (MGE: 1.2–1.6; RMSE: 1.6–2.0) and 4 km (MGE: 1.0–1.7; RMSE: 1.2–2.0). This indicates that the obtained modeling outcomes are characterized by a comparable scatter in relation to the mean observed wind speed. Simulated wind speed predictions are visibly biased. In the case of the site-specific (ONS) measurements, overestimation of simulated wind speed in relation to the observations at the level of 0.9-1.4 m/s was found for the WRF1 variant. An analogous relationship was noticed for the WRF4 variant as well, but the range of values is wider (from 0.3 to 1.5 m/s). In other cases, an underestimation of wind speed at the level of -1.1 to -0.1 m/s was observed (weather airport stations). A lower scatter of error values between experiments for the resolution of 1 km indicates that the use of higher resolutions of orography representation results in more stable (with a narrower range of error values) wind speed predictions. The achieved values of MB and IOA are not conclusive in terms of whether the increase in grid resolution results in the improvement of wind speed predictions. On the other hand, values of MCM shown in Figure 3, together with the confidence interval at the confidence level of 95%, indicate that apart from the TR (RTIA) experiment the obtained differences between variants are statistically significant (the confidence interval does not coincide with zero) for wind speed. This suggests that the grid resolution has a statistically significant impact on the outcomes of simulation predictions for wind speed in specific situations.



MED - MC - LV - TR

**Figure 3.** Graphical representation of the MCM values with confidence intervals at the 95% confidence level of the WRF model in the comparison of WRF accuracy evaluation for different grid resolutions (WRF1—WRF4), MED experiments and meteorological parameters (wind speed, wind direction, temperature) for synoptic stations (ISD).

In the case of the SIA and LVIA stations significantly lower values of mean gross error (MGE) were obtained for the computational grid resolution of 1 km compared to 4 km. However, for the WCA station the opposite observations were made. This is interesting due to the fact that in the LV experiment for two different locations (WCA and SIA) distinct results were obtained. The WCA station, located by the Westchester County Airport, is situated less than 10 km from the shoreline, which may cause discrepancies in the obtained evaluation results for different locations. Coastal regions are strongly affected by atmospheric forcing [96] that may not be accurately represented by numerical models if no additional, more accurate input data (e.g., sea surface temperature) are included in calculations [97,98] or the applied resolution is not sufficient to resolve terrain complexity and small-scale weather phenomena [84,99].

Increase in the computational grid resolution lowered the prediction bias, but simultaneously caused an increase in the absolute values of deviations (MGE). As a consequence, mean long-term predictions reflected well the actual conditions in the case of the resolution of 1 km, but this was not necessarily the case for short-term, e.g., 1 h, predictions. This finding is crucial in the context of air pollutant dispersion modeling for regulatory purposes. Achieving similar mean values for a simulation period (e.g., one year) indicates that a sample of results was obtained that will be statistically similar to reality, but will not always correctly represent momentary values. This is also a certain problem for air quality models used for regulatory purposes, which are mainly used to predict 1-h maximum and annual averaged values. When assessing the accuracy of predictions of air pollutant concentration levels, pairing of samples in time and space is generally neglected due to the imperfect nature of the applied methods. Awareness of these facts leads to a simple conclusion. In the evaluation of meteorological predictions carried out for air quality modeling, most likely it will be more critical to pay attention to the similarity of the prediction distributions compared to the observations. Thus, it is more important to obtain a sample of predictions, whose statistical characteristics will reflect the actual conditions. On the other hand, the disadvantage of such approach is the need to run simulations for a period of at least several years, in order to obtain a sample of data characterized by a full range of possible variations of meteorological conditions.

**Table 5.** Model evaluation measures for two computational grid resolutions with regard to the meteorological parameters at the ground surface and measurement location.

Parameters	Station	n	Mean		WRF1			WRF4	
	Station	п	Obs.	MB	MGE	RMSE	MB	MGE	RMSE
	MC-LVIA	8579	4.0	-0.7	1.3	1.8	-0.4	1.4	1.8
	MC-ONS	7835	2.4	1.1	1.5	1.9	1.4	1.7	2.0
Wind	LV-WCA	7253	3.9	-0.1	1.2	1.6	-0.3	1.1	1.5
wind	LV-SIA	5215	4.4	-0.4	1.5	2.0	-1.1	1.6	2.1
speed	LV-ONS	8647	1.2	1.4	1.6	2.0	1.5	1.6	2.0
	TR-RTIA	504	3.5	-1.1	1.6	2.0	-1.0	1.6	2.0
	TR-ONS	128	1.7	0.9	1.3	1.6	0.3	1.0	1.2
	MC-LVIA	8579	-	8.9	39.5	56.0	6.5	39.9	56.4
	MC-ONS	7835	-	-4.2	42.3	61.1	-4.1	41.8	59.0
147. 1	LV-WCA	7253	-	3.0	36.7	51.0	2.2	36.5	50.6
Wind	LV-SIA	5215	-	-6.6	35.4	49.3	-6.8	34.6	48.6
direction	LV-ONS	8632	-	0.9	60.0	77.6	-0.7	59.1	75.7
	TR-RTIA	504	-	-8.5	85.8	100.0	-8.4	84.9	98.7
	TR-ONS	128	-	19.0	72.6	92.4	6.5	76.3	95.4
	MC-LVIA	9192	10.8	-0.8	2.1	2.7	-2.4	3.0	3.8
Temperature	MC-ONS	9181	10.0	0.2	2.0	2.6	-1.6	2.5	3.2
	LV-WCA	8301	10.5	-1.1	2.3	3.1	-1.2	2.4	3.2
	e LV-SIA	7304	10.7	-1.9	3.0	3.8	-2.1	3.1	4.0
	LV-ONS	8698	10.3	-0.2	2.4	3.3	-0.3	2.6	3.6
	TR-RTIA	528	21.7	1.4	3.1	3.7	0.4	2.8	3.4
	TR-ONS	128	17.8	1.8	2.9	3.7	1.7	2.9	3.7

The results of the wind speed prediction assessment for the vertical profile of the atmosphere at each location are shown in Figure 4. In the case of the MC and LV experiments, significantly lower values of MGE for wind speed and 1 km resolution (WRF1) were obtained. In contrast, results for the TR experiment indicate that the obtained accuracy assessment results were not statistically significantly different. A wide confidence interval for the TR experiment is a consequence of the sample size (Supplementary Materials) and the accuracy of the WRF model is dependent on the season of the year [84]. Additional analyses, included in Supplementary Materials, indicate that when using the same data sample (the same time range of the year), significantly wider confidence intervals were obtained for the MC and LV experiments compared to those shown in Figures 3–5. Moreover, statistical model evaluation measures were generally higher compared to the results obtained for the entire period of the MC and LV experiments. Hence, each of the considered experiments (MC, LV and TR) should be treated independently.



MED - MC - LV - TR

**Figure 4.** Graphical representation of the MCM values with confidence intervals at the 95% confidence level in the comparison of the WRF accuracy evaluation for different grid resolutions (WRF1—WRF4), MED experiments and meteorological parameters (wind speed, wind direction, temperature) for on-site meteorological stations (ONS).

The above observations are important in the contest of the aim of the study, since increasing grid resolution generally results in a statistically significant increase in the accuracy of wind speed predictions at the on-site locations (in the case of datasets collected for at least one year). This in turn suggests that the use of meteorological predictions with a resolution of 1 km for the purpose of air quality modeling will provide wind speeds characterized by lower deviations from the actual values compared to coarser resolutions.

The presented results of the statistical parameters (MB, MGE, RMSE) of the wind direction assessment do not provide evidence that the increase in the computational grid resolution yields more reliable predictions (Table 5). The values of MGE and RMSE are similar at the resolutions of 1 km and 4 km. Values of MCM shown in Figure 3 demonstrate that only in the case of LV-SIA was a statistically significant difference between variants obtained. In other cases, the confidence intervals coincide with zero, indicating that the increase in the resolution of the computational grid did not result in an increase in the wind prediction accuracy. Values of MCM in the vertical profile (Figure 4) overwhelmingly indicate that there are no significant differences between the results obtained using 1 km and 4 km resolution. In the case of the MC experiment, different results were obtained for the comparative evaluation of different computational grids depending on the altitude above ground level. At the heights of 10 and 120 m negative values of MCM (-2.61 to -0.47) and for the heights of 300, 330 and 360 m positive values (0.02–0.97) were achieved. At the remaining heights above ground level the values of the MCM confidence interval coincide with zero. In conclusion, the obtained outcomes of the comparative evaluation of the WRF model's accuracy indicate that the use of 1 km grid resolution does not result in more reliable wind direction predictions. On the other hand, it is impossible to conclude that the resolution of 4 km yielded statistically significantly lower error values. However, in this case, the quality of predictions depends on the experiment or locations. A similar conclusion was presented by Giovannini et al. [50]. They pointed out that the increase in

the computational grid from a resolution of 2 km to 1.2 km resulted in higher accuracy of prediction only in selected locations (approximately 50% of the analyzed sites). The studies carried out by Giovannini et al. [50] refer to only one geographical area (Europe, Alps). They did not include the analysis of the accuracy of the WRF model in relation to the vertical profile, but measurements of meteorological parameters were made at 16 stations.

The obtained results of the WRF4 accuracy evaluation in the context of temperature (Table 5) indicate that, regardless of the grid resolution, the MGE values were greater than 2 for almost all cases. Similarly, for MB in six out of eight cases the values exceeded the range of |MB| < 0.5. Therefore, the WRF evaluation results do not meet the benchmarks proposed by [81]. However, in the case of the MC and LV experiments, the index of agreement reached IOA  $\geq 0.8$  (see Supplementary Materials), which is in agreement with the criteria established in [81]. It should be emphasized that the presented results of the temperature prediction accuracy do not differ from the outcomes in other studies concerning computational grid resolution [1,63,86,100]. Comparative analysis of the temperature modeling evaluation unambiguously indicates that the use of a finer grid resolution (1 km) yields higher accuracy. Both in Figures 3 and 4, the MCM parameter with confidence intervals at the confidence level of 95% has positive values for the MC and LV experiments. In contrast, the results for the TR experiment are inconclusive. Statistically significant and lower errors for temperature predictions were obtained only for the altitudes above 250 m a.g.l. In other cases, generally no statistically significant differences were found (Figure 4). The presented conclusions are in agreement with the results of studies [63,100] where a similar increase in temperature prediction accuracy was identified with the use of finer grid resolutions. However, [100] pointed out that the resolution of 3 km is satisfying, since lesser computational costs allowed more simulations to be carried out in the same amount of time. On the other hand, it should be emphasized that in the aforementioned studies the confidence intervals of the model evaluation measures were not determined, and thereby no statistically significant differences between evaluated configurations were estimated. Therefore, a direct comparison of the results between studies is impossible. In addition, the study above did not include an accuracy assessment of the WRF model along the vertical profile of the atmosphere.

Taking into account the obtained results, we recommend using a resolution of 1 km, because it allows the achieving of statistically significant and lower MGE values compared to a grid with a resolution of 4 km. Therefore, the remaining Sections 3.2 and 3.3 are limited to the results only for the resolution of 1 km. The outcomes presented above and in the discussion indicate that the use of 1 km resolution compared to 4 km usually increases the accuracy of wind speed and temperature predictions. The results for wind direction were ambiguous but no contradictions to the above conclusion were found. Future research on grid resolution should focus on parameterizing the model by geographic region and number of pressure levels to see if boundary layer issues in complex terrain are correctly resolved. Better representation of topography will not increase the accuracy of numerical model predictions due to the inability of numerical models to represent near-surface structures, e.g., temperature inversions [101]. This problem is very well illustrated by the animation of vertical temperature profiles presented in the Supplementary Materials (Chapter S4).

#### 3.2. Location of the Station

A comparison of the results was conducted for two types of the stations, i.e., on-site (ONS) and synoptic (ISD) situated in different locations. The analysis was aimed at verifying whether the accuracy of WRF model predictions is similar in non-typical locations compared to standard meteorological measurements at sites located near airports. Unlike many other studies [84,85,102], the presented evaluation includes campaigns datasets from the MC, LV and TR experiments. Measurements were carried out at various levels relative to the ground surface. In addition, these data were not assimilated to the simulations at any stage. In effect, the evaluation of meteorological predictions was possible in areas where standard measurement data are not available. This situation corresponds to the realities of air quality

modeling of pollutant dispersion in the atmosphere, where in situ meteorological data are not always available. In the case of prediction accuracy comparison for two different locations it was necessary to unify the frequency of observations between data from stations located near airports (ISD) and data obtained as part of experiments (ONS). MB, MGE and RMSE results with confidence intervals at the 95% confidence level, calculated for temperature, wind speed and direction for each experiment, are shown in Figure 5.



**Figure 5.** Graphical representation of the MB, MGE and RMSE model performance measures for the WRF evaluation divided into meteorological parameters, experiments and the location of the measurement site inside computational domain.

The MGE values obtained for wind speed are closer to zero for ISD stations compared to ONS sites. The 95% confidence intervals did not overlap for the MC and LV experiments. This proves that the achieved WRF accuracy results at different locations within the same computational domain are statistically significantly different. The MB value explicitly demonstrates that the WRF model was biased with respect to the station location. For the ONS sites the results were overestimated, which is well illustrated by the scatter plots in Supplementary Materials (Chapter S2.2.1). Similar conclusions regarding the WRF model bias were presented in [101,103]. In the case of ISD stations, wind speeds were underestimated in the range from -1.12 m/s to -0.12 m/s. The RMSE values indicated that the scatter of the results for the MC (1.83 m/s-1.86 m/s) and TR (1.48 m/s-1.62 m/s) experiments was similar for different sites (95% confidence intervals overlapped) and for both comparable cases in the LV experiment it was different (ranging from 1.5 m/s to 2.06 m/s). Wind speed deviations were high depending on the site location. It is manifested by the MB values which, in some ONS cases, were at the level of mean observed values. According to Hanna and Yang [104] the uncertainty of wind speed stems majorly from random turbulent processes as well as errors in both land use and topography representation. Taking into account the outcomes for ONS and ISD stations it is suspected that in this study the main component of the prediction uncertainty was the location of the ONS stations in complex terrain. It is suspected that the problem of overestimation of wind speed in complex terrain can be corrected by using the parameterization of the WRF model described in Jiménez and Dudhia [68]. In the future, parameterization of the WRF model should be considered in order to minimize wind speed errors in complex terrain (on-site station locations).

The presented results raise doubts in the context of the use of wind speed prediction for air quality modeling. The use of the overestimated predictions in the air quality model will result in lower air pollutant concentrations compared to the observed values and, what is more, the maximum concentrations will be located at a greater distance from the emission source. In the case of the air quality model, additional attention should be drawn to the plume rise. It is dependent on the wind speed and temperature difference between the plume and the environment, as well as the plume ejection velocity. For wind speed that dependence is inversely proportional [105]. In effect, the wind speed overprediction in the WRF model will result in a lower plume rise in the considered experiments.

All of the considered model performance measures for wind direction in almost every case were statistically significantly different (Figure 5). This parameter was the most sensitive to the representation of actual conditions in different locations. Therefore, for this parameter, the values of wind direction deviations are presented using polar plots in Figure 6. Graphical representation indicates unambiguously that in the case of the ONS sites deviations exceeding  $\pm 90^{\circ}$  are more frequent. This is reflected by the statistical model performance measures. The MGE values for ISD locations were at the level of 40° but for the ONS sites they were significantly higher and exceeded the value of  $60^{\circ}$  (LV, MC). Similar observations were noted for RMSE (ISD—RMSE < 60°, ONS—RMSE > 60°). The above remarks are also confirmed by the comparative analysis of wind roses included in Supplementary Materials. For example, in the MC experiment, due to the location of the ONS station in a relatively narrow valley of the Delaware River (with a width not greater than 15 km in this region), two dominant, opposite wind directions are noticeable—SW and NE. These directions were partially reproduced by the model as manifested by MGE and RMSE values of  $42^{\circ}$  and  $61^{\circ}$ , respectively. However, due to the lack of natural obstacles in the vicinity of the airport, the wind rose for the ISD (LVIA) station does not show the presence of dominant directions of air inflow, only a typical predominance of winds blowing from the west. This situation was represented more accurately by the model simulations and resulted in MGE and RMSE values for wind direction at the level of  $36^\circ$ and  $51^{\circ}$ , respectively. Similar conclusions were shown in [106], where it was observed that for a location with complex terrain the RMSE values of wind direction were twice as high compared to areas located on flat terrain.



**Figure 6.** Graphical representation of wind direction deviations at the surface (10 m a.g.l.) depending on the type of the station (ISD, ONS) and experiment (MC, LV, TR).

Liu et al. [106] unlike this study, used the assimilation of observational data, and yet they did not obtain an effective improvement in the results of wind direction forecasts in complex terrain. As in Jiménez and Dudhia [80], it was observed that wind direction errors are inversely related to wind speed. This is presented in the graphs of MGE and RMSE values depending on wind speed, included in Supplementary Materials (Chapters S2.2 and S3.2.2). This effect is a consequence of topography, which strongly modifies circulation at the synoptic scale. This results in a large spatial variability of the airflow in complex terrain. The reason for these errors is the limitations of the WRF model's representation of topography, as synoptic circulation is generally well represented by mesoscale models [80]. In contrast to these studies, they used a horizontal grid resolution of 2 km, and the accuracy assessment referred only to wind direction measurements at a height of 10 m a.g.l. Therefore, it can be concluded that reducing the horizontal grid resolution to 1 km has not solved this problem in complex terrain. However, in contrast to previous studies, the presented results of the WRF model accuracy assessment (Supplementary Materials—Chapter S3.2.2), provide evidence that the problem is not only related to the representation of wind direction at a height of 10 m a.g.l. It affected all vertical layers for which the accuracy of the WRF model was evaluated. It is suspected to be the consequence of greater variability in wind direction at low wind speeds, which is more difficult to represent in meteorological simulations.

In the case of the TR experiment both MB and RMSE proved the opposite observations. Lower error values were noted for the ONS station. On the other hand, the RTIA site is situated in Reno which is located on the border of the Mount Rose wilderness. Moreover, the RTIA measurement site is located relatively close to the computational domain border in conducted simulations (about 4.5 km away). Complex topography and location of the station might be the reasons for higher deviations of wind direction, especially considering that the remaining ISD stations (SIA, WCA, LVIA) are located in areas characterized by a less varied topography in terms of the absolute altitude. Extending the size of the domain might address this issue and enable dynamics and physics to fully evolve inside and correct possibly improper large-scale features inherited from the lateral boundary conditions [107,108].

Application of the WRF meteorological model data in air pollutant dispersion modeling will result in specific consequences. Atmospheric wind direction is equivalent to the direction of the plume from the emission source. In effect the air pollutant concentration levels obtained from the air dispersion model can be observed in completely different areas. Measurement of the error of wind direction at the level of  $5^{\circ}-10^{\circ}$  results in calculated concentration uncertainty ranging from 20% to 70% [38]. The wind roses included in Supplementary Materials show that the contributions of the main wind directions were represented properly in experiments TR and LV. Theoretically, it can be presumed that a sufficient criterion for air quality modeling may be the appropriate representation of the wind directions. This results from the methodology for the air quality model evaluation for the purposes of impact assessment of the source on air quality. Suitable representation of the upper distributions of the 26 highest 1-h concentrations using the RHC (Robust Highest Concentration) parameter is the most important [109,110]. On the other hand, the overlapping of errors originating from meteorological simulations and determination of air pollutant concentrations can create a "domino effect". This might lead to unrepresentative modeling results, significantly deviating from the actual values. In addition, an opposite effect may be achieved, which leads to more accurate air quality predictions, as a result of unintended compensation of errors. Both of these scenarios are plausible, hence practical application of these techniques must allow for certain risks. Determination of the origins of discrepancies requires further research comparing the accuracy of air quality predictions from the analyzed experiments with the application of various input meteorological data. However, the most direct proof of the possibility of error compensation is the results of research prepared for the MC experiment by US. EPA [6]. Comparison of the wind roses obtained in this study and those presented in US. EPA [6] implies that the model failed to adequately represent dominant wind directions. The wind speed was biased as well, as evidenced by the presented MB = 0.59 m/s. However, this value is still lower than the one obtained in this study (MB = 1.11 m/s). It is suspected that the reason for the meteorological prediction discrepancy is associated with the WRF domains set up. Grid cells did not overlap which could have resulted in different outcomes due to the terrain elevation. Despite the fact that the accuracy of the meteorological simulations differed from the actual conditions, the AERMOD evaluation results presented in US. EPA [6] did not differ statistically significantly compared to simulations using measurement data from the weather tower. Moreover, a number of studies have successfully applied meteorological simulation data to model pollutant concentration levels in the air [21,23,35,42–47]. Accuracy of the used meteorological simulations differed as well from the actual conditions. Therefore, it is suspected that the main reason for similar AERMOD evaluation results for different meteorological input data is a consequence of disparate methodologies for model evaluation and unintended compensation of the output data.

Evaluation results for temperature demonstrated that the MB values were closer to zero for on-site stations compared to weather stations (Figure 5), in contrast to the wind speed and direction evaluation. Achieved MB values together with 95% confidence intervals did not overlap. Therefore, the WRF evaluation outcomes with regard to the location are statistically significantly different. Similar conclusions can be drawn from the MGE analysis except for the LV-WCA case. On the other hand, RMSE values for different experiments and locations within a given experiment are not conclusive. In the case of the LV experiment, opposite outcomes were obtained depending on the locations (LV-SIA vs. LV-WCA). It is worth highlighting that, in the case of the TR experiment, the MB values are positive, whereas for the other experiments negative values were obtained. Data from the TR experiment were collected only in the summer period (August 1984) which implies that the WRF model is sensitive to seasonality. This can be observed on the "Statistical metrics" graphs presented in Supplementary Materials, where, depending on the temperature range, the MB values were different. In general, for temperatures lower than 15  $^{\circ}$ C, the MB values were lower than zero (underestimation of observations). For higher temperatures MB values were greater than zero (overestimation of observations). The presented outcomes

are atypical in the case of the WRF model, which generally struggles with representation of the highest temperatures, especially for long-term simulations [16]. On the other hand, the TR experiment was conducted in 1984. For this study, the NLCD 1992 land use data was used, which most likely was characterized by a higher share of urban tissue and built-up area compared to the actual conditions. The presence of urban and built-up areas affects simulated sensible and latent heat fluxes causing an increase in air temperature which may lead to overestimation by the meteorological model [65].

#### 3.3. Vertical Profile of the Atmosphere

In the case of the wind speed evaluation, it was decided to apply normalized evaluation measures (NMB, NMGE) since they are resistant to data with different ranges of values. For instance, wind speed values at the highest measurement points were up to 3–4 times greater than the values at 10 m a.g.l. Comparison using basic statistics, i.e., MB, MGE, RMSE could distort the interpretation and lead to incorrect observations. Error values presented in Figure 7 indicate that the meteorological simulation accuracy depends on the altitude above ground level as well. As the height above ground level increased for the MC, LV and TR experiments, the NMGE values decreased. In the case of the MC experiment, they decreased from 0.78 to 0.39 (10 m-420 m a.g.l.) and for the LV the difference was more apparent from 1.27 to 0.74 with a change in altitude from 10 m to 100 m a.g.l. For the TR experiment, NMGE decreased from 0.77 to 0.48 (10 m–400 m a.g.l.). It should be emphasized that the greatest differences between the model evaluation outcomes (statistically significantly different) for particular altitudes above terrain occur up to 120 m about ground level. For wind direction (Figure 8) there is a similar dependency; however, it applies to the lowest layer. At higher altitudes differences between given layers are generally lower and often statistically insignificant, as proven by not overlapping 95% confidence intervals (Figures 7 and 8). The presented outcomes unambiguously indicate that the representation of local conditions, where complex terrain issue is of great importance, constitutes a major challenge. Moreover, it is strongly location-dependent, since for the MC and LV experiments statistically significantly different values of NMBE were achieved (respectively, 0.62 and 1.62 at the surface layer). This presents the problem of applying the meteorological simulation data in air quality models, because, depending on the emission source (or group of sources) location, the quality of input data in the form of meteorological predictions may be statistically significantly different. Since a more accurate representation of topographic characteristics does not result in a clear improvement in performance [50,101], future studies should also consider increasing the resolution of vertical layers.

The obtained meteorological predictions of wind direction were biased regardless of the experiment (Figure 8). The positive values of MB indicate that the meteorological predictions were mostly right-sided compared to the observations. The values of absolute errors (MGE) were 2–3 times greater than the MB values. This implies that there are both left- and right-sided deviations of wind direction, on average considerably higher than indicated by the MB value. Included in Supplementary Materials the "wind direction error" and "wind direction error histogram" graphs point out that the deviations in the range from  $-45^{\circ}$  to  $60^{\circ}$  were dominant for the MC and LV experiments. In the case of the TR experiment the histograms were bimodal, with two dominant ranges of deviations present, i.e., from  $-180^{\circ}$  to  $-90^{\circ}$  and from  $-20^{\circ}$  to  $90^{\circ}$ . It should be noted that starting from the 100 m a.g.l. the accuracy of the model was similar in the case of the MC experiment and for the TR experiment there was a noticeable improvement in the prediction accuracy with altitude. An important finding is that regardless of altitude, meteorological simulations of wind direction have an error that is inversely proportional to wind speed in complex terrain (Supplementary Materials—Chapter S3.2.2). As wind speed increases, the MGE and RMSE values approach zero, regardless of the height above ground level. This is a known problem of the WRF model [80], but its existence has so far been proven for the near-surface layer (heights of about 10 m a.g.l.).



MED - MC - LV - TR

**Figure 7.** Graphical representation of the NMB, NMGE and IOA values calculated for wind speed with regard to the altitude above ground level and considered experiment.



#### MED 🔶 MC 🛶 LV 🛶 TR

**Figure 8.** Graphical representation of the MB, MGE and RMSE values calculated for wind direction with regard to the altitude above ground level and considered experiment.

The aspect of wind speed and direction prediction accuracy along the vertical profile is important for the purpose of air quality modeling. Emission sources diverge in terms of, among others, the height at which gases and particulate matter are introduced into the atmospheric air. In the case of point sources, the height of emitters reaches from several to 200 m, and the height of most power industry fossil fuel combustion emitters often exceeds 300 m, and in extreme cases even 400 m (Ekibastuz GRES-2 Power Station). In air pollutant dispersion models, the determination of wind speed and direction on the effective height of the emitter is crucial. This value affects the process of transport and diffusion of air pollutants. In the absence of measurement data, wind speed is estimated using different

meteorological preprocessors, such as AERMET or CALMET. The methodology based on the Monin–Obukhov (M–O) similarity theory for wind profile determination depending on the atmospheric stratification is frequently applied. It should be emphasized that the involved method constitutes only an approximation. The greatest discrepancies occur for stable stratification conditions due to the formation of intermittent layers and for complex terrain areas [111,112]. Therefore, in such situations, using data from meteorological models which provide additional information about vertical wind speed and direction profiles, appears reasonable. Considering the presented results of the WRF prediction evaluation, in terms of wind speed and direction along the vertical profile, a certain issue emerges. Application of meteorological predictions in air quality modeling for a group of emitters characterized by various heights (e.g., 40 m, 100 m, 200 m) may result in entirely different errors of pollutant concentration levels due to the incorrect representation of meteorological conditions along the vertical profile of the atmosphere. In addition, the location of the emission source impact area may also be oriented in other directions.

The obtained results of the WRF evaluation in the case of NMB unambiguously indicate that regardless of the experiment, the model was biased. For MC and LV experiments the overestimation of wind speed was found (MC: NMB = 0.19-0.62; LV: NMB = 0.54-1.13), which decreased with the increase in altitude. In many cases the differences between individual vertical layers were statistically significantly different. In general, the greatest difference of NMB, NMGE and IOA values between altitudes were obtained for the heights from 10 to 120 m. This implies that the WRF model is not able to represent the actual meteorological conditions in locations where topography characteristics vary within one cell of computational grid. This is proven by the comparison of NMB, NMGE, IOA (Figure 7), MB, MGE, RMSE (Figure 8) and wind roses included in Supplementary Materials (Chapter S3.1.2) obtained for the LV experiment. WRF predictions for wind direction in the LV experiment indicate that the winds blowing from the west were dominant with the share of 22%. The values recorded by the on-site station show that the actual contribution of western winds was only 6%. Moreover, the WRF model failed to appropriately represent winds blowing from the northern and south-eastern directions, i.e., winds along the Hudson River valley. Based on the analysis of actual topography compared to the applied digital terrain model with 1 km accuracy it can be observed that there was a clear flattening of Buckberg Mountain. The peak of this hill was located approximately 600–700 m from the meteorological tower and both of these points were mapped within one cell of the computational grid. In consequence the representation of terrain blocking effects was impossible, which resulted in reduction in the western winds. It is not completely understandable why the WRF model failed to accurately represent the contribution of northern and south-eastern winds. The conducted discussion of this issue leads to a conclusion. In the case of meteorological prediction application (when measurement data are unavailable) in the air quality modeling process, first, the expert analysis of wind roses should be carried out in the locations of considered emission sources with regard to the topography. In many cases it would be possible at the expert level to determine the potential of using meteorological predictions.

In Figure 9 the values of NMB, NMGE and IOA calculated for temperature values with regard to the altitude above ground level are presented. NMB and NMGE were, for most cases, statistically significantly different depending on the altitude above ground. These differences are particularly clear for the values at the ground surface. IOA values for the MC and LV experiments were at the levels from 0.86 to 0.87 and met the criteria recommended by [81]. In the case of the TR experiment the values of IOA were lower and, depending on the altitude above ground level, fell within the range of 0.44 to 0.62. On the other hand, the NMGE values were statistically significantly lower compared to the values obtained for the MC and LV experiments. In order to better understand these issues, the animations showing vertical profiles of temperature for the TR experiment are presented in Supplementary Materials. In the analyzed period, an increase in temperature with altitude (temperature inversions) prevailed—most often at heights up to 200 m a.g.l. Modeling

results generally misrepresented the typical vertical temperature profile. In most situations, the temperature decreased with the increase in altitude. The values of the temperature gradient are commonly used for atmospheric stability determination, which affects both the horizontal and vertical diffusion coefficients applied in the air quality models.



MED 🔶 MC 🔶 LV 🛶 TR

**Figure 9.** Graphical representation of the NMB, NMGE and IOA values calculated for temperature with regard to the altitude above ground level and experiment.

## 4. Conclusions

The study was carried out concerning the evaluation of the WRF meteorological simulations (temperature, wind speed and direction), with regard to the computational grid resolution, location in complex terrain and the vertical profile of the atmosphere. The conducted research allowed for the formulation of several conclusions in the context of the application of meteorological predictions in air quality models.

Increasing the grid resolution to 1 km did not always result in statistically significantly lower error values compared to a resolution of 4 km. In most cases wind roses at 4 km did not represent the dominant wind directions, the same as for the on-site station in the LV experiment at 1 km. Hence, an increase in the resolution up to 1 km will not always assure adequate accuracy of the input data for air pollutant dispersion models. Therefore, this aspect should be treated individually in relation to the orography of the area. On the other hand, a further increase in the resolution is not recommended due to the suggestions presented in [113]. It is noteworthy that the optimization of the WRF simulation setup [114,115] and the application of model coupling [116–118] proved to be effective to further improve the accuracy of modeling outcomes.

A dependency between locations of the meteorological measurements and the outcomes of meteorological simulation accuracy was observed for wind speed and temperature. In the case of on-site stations (complex terrain) greater error values were obtained compared to the standard ISD location (airport). Similar issues were reported in [50,84]. Therefore, depending on the location of the emission source within the computational domain, the accuracy of air pollutant dispersion simulations may be different. This issue is crucial in the case of applications where emission sources are located in complex terrain and nearby representative measurement data are unavailable.

The evaluation of the WRF meteorological simulations in the vertical profile of the atmosphere with regard to wind speed and direction in the on-site locations implies that the accuracy changes are statistically significantly at the confidence level of 95%. This refers

to the vertical profile up to the altitude of 100 m a.g.l. Generally, the accuracy of predictions increases with altitude, which is most likely the consequence of decreasing effect of terrain orography. Above the height of 100 m a.g.l. the prediction accuracy stabilizes at a given level or increases further with the altitude, depending on the experiment. The increase in accuracy with altitude has certain consequences for the diffusion of air pollutants. Depending on the type of emission sources, with a particular attention to the height at which gases and particulate matter are introduced into the atmospheric air, the uncertainty of these simulations will depend on the prediction accuracy at different altitudes.

Clear positive bias of the meteorological predictions in areas characterized by complex terrain in relation to the observed wind speed values, in the case of application in air quality models for regulatory purposes, may result in obtaining significantly lower values of concentration levels of the considered air pollutants compared to the actual values. In effect, this may lead to severe consequences during the implementation of recovery programs for air quality, e.g., qualifying a given source or group of sources as negligible based on false evidence. It is possible that the use of the WRF model parameterization described by Jiménez and Dudhia [68,80] will allow the elimination of these effects.

Considering the accuracy of meteorological predictions carried out using the WRF model, their use in air quality modeling is recommend only in certain circumstances. Firstly, when the measurement data are not available and their acquisition is impossible or economically unjustified. The above is mostly in line with the recommendations [38]. Moreover, for each case study, an evaluation of the meteorological prediction accuracy should be conducted with regard to the orography features in the vicinity of emission sources, in particular when the study area is characterized by complex terrain. It should also be emphasized that in the case of lack of measurement on-site data, the evaluation of the WRF model using information from the ISD database might not be sufficient. In addition, it is advised that in areas planned for investments that could potentially significantly affect air quality, meteorological conditions should be measured before issuing an environmental decision.

It has been demonstrated that there is a risk of an unintended compensation of the uncertainty of air quality modeling outcomes [6] or lowering the accuracy of predictions by introduction of inaccurate meteorological data [2,5,48]. It has been stated that this might result from differences in methodologies for air quality and meteorological model evaluation.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/su151612576/s1. Supplementary Materials.html.

**Author Contributions:** Conceptualization, M.R., A.K. and R.O.; methodology, M.R. and A.K.; software, M.R. and A.K.; validation, M.R.; formal analysis, M.R. and A.K.; investigation, M.R. and A.K.; resources, M.R. and A.K.; data curation, M.R. and A.K.; writing—original draft preparation, M.R. and A.K.; writing—review and editing, M.R. and R.O.; visualization, M.R.; supervision, M.R.; project administration, M.R.; funding acquisition, M.R. and R.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was carried out as part of research associated with the Ministry of Education and Science (Poland) subsidy to maintain scientific potential (Contract no. 16.16.150.545).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The additional data presented in this study are available in Supplementary Materials. Other data sharing is not applicable to this article. It is possible to provide modeling results on request.

Conflicts of Interest: The authors declare no conflict of interest.

## References

- Tartakovsky, D.; Stern, E.; Broday, D.M. Evaluation of modeled wind field for dispersion modeling. *Atmos. Res.* 2015, 166, 150–156.
   [CrossRef]
- 2. Abril, G.A.; Diez, S.C.; Pignata, M.L.; Britch, J. Particulate matter concentrations originating from industrial and urban sources: Validation of atmospheric dispersion modeling results. *Atmos. Pollut. Res.* **2016**, *7*, 180–189. [CrossRef]
- 3. Karthick, J.; Devadoss, R.S. Behaviour and the influence of meteorological parameters towards air pollutant dispersion modelling studies, Manali industrial area, Chennai, India. *Ecol. Environ. Conserv.* **2018**, 24, 1822–1828.
- 4. Ottosen, T.B.; Ketzel, M.; Skov, H.; Hertel, O.; Brandt, J.; Kakosimos, K.E. Micro-scale modelling of the urban wind speed for air pollution applications. *Sci. Rep.* 2019, *9*, 14279. [CrossRef]
- Rzeszutek, M.; Szulecka, A.; Oleniacz, R.; Bogacki, M. Assessment of the AERMOD dispersion model over complex terrain with different types of meteorological data: Tracy Power Plant experiment. *E3S Web Conf.* 2017, 22, 00149. [CrossRef]
- U.S. EPA. Evaluation of Prognostic Meteorological Data in AERMOD Applications; EPA-454/R-18-002, U.S. EPA RTP, NC.; U.S. EPA: Washington, DC, USA, 2018.
- Sistla, G.; Zhou, N.; Hao, W.; Ku, J.Y.; Rao, S.T.; Bornstein, R.; Freedman, F.; Thunis, P. Effects of uncertainties in meteorological inputs on urban airshed model predictions and ozone control strategies. *Atmos. Environ.* 1996, 30, 2011–2025. [CrossRef]
- 8. Gilliam, R.C.; Hogrefe, C.; Rao, S.T. New methods for evaluating meteorological models used in air quality applications. *Atmos. Environ.* **2006**, *40*, 5073–5086. [CrossRef]
- 9. Pielke, R.A.; Uliasz, M. Use of meteorological models as input to regional and mesoscale air quality models—Limitations and strengths. *Atmos. Environ.* **1998**, *32*, 1455–1466. [CrossRef]
- 10. McNider, R.T.; Pour-Biazar, A. Meteorological modeling relevant to mesoscale and regional air quality applications: A review. J. Air Waste Manag. Assoc. 2019, 70, 2–43. [CrossRef]
- Skamarock, W.C.; Klemp, J.B.; Dudhi, J.; Gill, D.O.; Barker, D.M.; Duda, M.G.; Huang, X.-Y.; Wang, W.; Powers, J.G. A Description of the Advanced Research WRF Version 3; NCAR/TN-475+STR.; University Corporation for Atmospheric Research: Boulder, CO, USA, 2008.
- 12. Grell, G.A.; Dudhia, J.; Stauffer, D.R. A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5); University Corporation for Atmospheric Research: Boulder, CO, USA, 1994.
- Gómez, I.; Caselles, V.; Estrela, M.J.; Miró, J.J. Comparative assessment of RAMS and WRF short-term forecasts over Eastern Iberian Peninsula using various in-situ observations, remote sensing products and uncoupled land surface model datasets. *Atmos. Res.* 2018, 213, 476–491. [CrossRef]
- 14. Giannaros, C.; Melas, D.; Giannaros, T.M. On the short-term simulation of heat waves in the Southeast Mediterranean: Sensitivity of the WRF model to various physics schemes. *Atmos. Res.* **2019**, *218*, 99–116. [CrossRef]
- 15. Perdigão, J.; Salgado, R.; Magarreiro, C.; Soares, P.M.M.; Costa, M.J.; Dasari, H.P. An Iberian climatology of solar radiation obtained from WRF regional climate simulations for 1950–2010 period. *Atmos. Res.* **2017**, *198*, 151–162. [CrossRef]
- 16. Marta-Almeida, M.; Teixeira, J.C.T.; Carvalho, M.J.; Melo-Gonçalves, P.; Rocha, A.M. High resolution WRF climatic simulations for the Iberian Peninsula: Model validation. *Phys. Chem. Earth* **2016**, *94*, 94–105. [CrossRef]
- 17. Mooney, P.A.; Mulligan, F.J.; Bruyère, C.L.; Parker, C.L.; Gill, D.O. Investigating the performance of coupled WRF-ROMS simulations of Hurricane Irene (2011) in a regional climate modeling framework. *Atmos. Res.* **2019**, 215, 57–74. [CrossRef]
- 18. Shen, F.; Min, J.; Xu, D. Assimilation of radar radial velocity data with the WRF Hybrid ETKF-3DVAR system for the prediction of Hurricane Ike (2008). *Atmos. Res.* **2016**, *169*, 127–138. [CrossRef]
- González-Alonso de Linaje, N.; Mattar, C.; Borvarán, D. Quantifying the wind energy potential differences using different WRF initial conditions on Mediterranean coast of Chile. *Energy* 2019, 188, 116027. [CrossRef]
- 20. Prósper, M.A.; Otero-Casal, C.; Fernández, F.C.; Miguez-Macho, G. Wind power forecasting for a real onshore wind farm on complex terrain using WRF high resolution simulations. *Renew. Energy* **2019**, *135*, 674–686. [CrossRef]
- 21. Abdul-Wahab, S.A.; Ali, S.; Sardar, S.; Irfan, N.; Al-Damkhin, A. Evaluating the performance of an integrated CALPUFF-MM5 modeling system for predicting SO<sub>2</sub> emission from a refinery. *Clean Technol. Environ. Policy* **2011**, *13*, 841–854. [CrossRef]
- 22. Gonzalez, J.A.; Hernandez-Garces, A.; Rodriguez, A.; Saavedra, S.; Casares, J.J. A comparison of different WRF-CALMET simulations against surface and PBL rawinsonde data. In Proceedings of the 16th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Varna, Bulgaria, 8–11 September 2014; Volume 137, pp. 52–56.
- Lee, H.-D.; Yoo, J.-W.; Kang, M.-K.; Kang, J.-S.; Jung, J.-H.; Oh, K.-J. Evaluation of concentrations and source contribution of PM10 and SO2 emitted from industrial complexes in Ulsan, Korea: Interfacing of the WRF-CALPUFF modeling tools. *Atmos. Pollut. Res.* 2014, 5, 664–676. [CrossRef]
- 24. Affum, H.A.; Akaho, E.H.K.; Niemela, J.J.; Armenio, V.; Danso, K.A. Validating the California Puff (CALPUFF) Modelling System Using an Industrial Area in Accra, Ghana as a Case Study. *Open J. Air Pollut.* **2016**, *5*, 27–36. [CrossRef]
- 25. Afzali, A.; Rashid, M.; Afzali, M.; Younesi, V. Prediction of air pollutants concentrations from multiple sources using AERMOD coupled with WRF prognostic model. *J. Clean. Prod.* 2017, *166*, 1216–1225. [CrossRef]
- Tao, Z.; Chin, M.; Gao, M.; Kucsera, T.; Kim, D.; Bian, H.; Kurokawa, J.-I.; Wang, Y.; Liu, Z.; Carmichael, G.R.; et al. Evaluation of NU-WRF model performance on air quality simulation under various model resolutions-an investigation within the framework of MICS-Asia Phase III. *Atmos. Chem. Phys.* 2020, 20, 2319–2339. [CrossRef]

- Fountoukis, C.; Mohieldeen, Y.; Pomares, L.; Gladich, I.; Siddique, A.; Skillern, A.; Ayoub, M.A. Assessment of High-resolution Local Emissions and Land-use in Air Quality Forecasting at an Urban, Coastal, Desert Environment. *Aerosol Air Qual. Res.* 2022, 22, 220001. [CrossRef]
- 28. Gadzhev, G.; Ganev, K.; Mukhtarov, P. Influence of the Grid Resolutions on the Computer-Simulated Surface Air Pollution Concentrations in Bulgaria. *Atmosphere* **2022**, *13*, 774. [CrossRef]
- Hinestroza-Ramirez, J.E.; Lopez-Restrepo, S.; Yarce Botero, A.; Segers, A.; Rendon-Perez, A.M.; Isaza-Cadavid, S.; Heemink, A.; Quintero, O.L. Improving Air Pollution Modelling in Complex Terrain with a Coupled WRF–LOTOS–EUROS Approach: A Case Study in Aburrá Valley, Colombia. *Atmosphere* 2023, 14, 738. [CrossRef]
- Soares da Silva, M.; Cláudio Gomes Pimentel, L.; Pereira Duda, F.; Aragão, L.; Silva, C.; Cunha, I.D.; Viana Dragaud, A.; Caffaro Vicentini, P. Assessment of meteorological settings on air quality modeling system-a proposal for UN-SDG and regulatory studies in non-homogeneous regions in Brazil. *Environ. Sci. Pollut. Res.* 2023, *30*, 1737–1760. [CrossRef] [PubMed]
- 31. Holnicki, P.; Kałuszko, A.; Trapp, W. An urban scale application and validation of the CALPUFF model. *Atmos. Pollut. Res.* **2016**, 7, 393–402. [CrossRef]
- 32. Holnicki, P.; Kałuszko, A.; Nahorski, Z.; Stankiewicz, K.; Trapp, W. Air quality modeling for Warsaw agglomeration. *Arch. Environ. Prot.* **2017**, *43*, 48–64. [CrossRef]
- Hao, Y.; Xie, S. Optimal redistribution of an urban air quality monitoring network using atmospheric dispersion model and genetic algorithm. *Atmos. Environ.* 2018, 177, 222–233. [CrossRef]
- Wu, H.; Zhang, Y.; Yu, Q.; Ma, W. Application of an integrated WRF/CALPUFF modeling tool for source apportionment of atmospheric pollutants for air quality management: A case study in the urban area of Benxi, China. *J. Air Waste Manag. Assoc.* 2017, 68, 347–368. [CrossRef] [PubMed]
- Kesarkar, A.P.; Dalvi, M.; Kaginalkar, A.; Ojha, A. Coupling of the Weather Research and Forecasting Model with AERMOD for pollutant dispersion modeling. A case study for PM10 dispersion over Pune, India. *Atmos. Environ.* 2007, 41, 1976–1988. [CrossRef]
- Boadh, R.; Satyanarayana, A.N.V.; Rama Krishna, T.V.B.P.S.; Madala, S. Sensitivity of PBL parameterization schemes of Weather Research Forecasting Model and coupling with AERMOD in the dispersion of NOX over Visakhapatnam (India). *Asia-Pac. J. Chem. Eng.* 2015, 10, 258–261. [CrossRef]
- 37. Kumar, A.; Patil, R.S.; Dikshit, A.K.; Kumar, R. Application of AERMOD for short-term air quality prediction with forecasted meteorology using WRF model. *Clean Technol. Environ. Policy* **2017**, *19*, 1955–1965. [CrossRef]
- U.S. EPA. Revisions to the Guideline on Air Quality Models: Enhancements to the AERMOD Dispersion Modeling System and Incorporation of Approaches to Address Ozone and Fine Particulate Matter; 40 CFR Part 51, Appendix W. Federal Register; U.S. EPA: Washington, DC, USA, 2017; Volume 82.
- Smith, A.; Lott, N.; Vose, R. The integrated surface database: Recent developments and partnerships. *Bull. Am. Meteorol. Soc.* 2011, 92, 704–708. [CrossRef]
- 40. Durre, I.; Vose, R.S.; Wuertz, D.B. Overview of the integrated global radiosonde archive. J. Clim. 2006, 19, 53–68. [CrossRef]
- 41. Rzeszutek, M.; Bogacki, M.; Paulina, B.; Szulecka, A. Improvement assessment of the OSPM model performance by considering the secondary road dust emissions. *Transp. Res. D Transp. Environ.* **2019**, *68*, 137–149. [CrossRef]
- Kumar, A.; Patil, R.S.; Dikshit, A.K.; Islam, S.; Kumar, R. Evaluation of control strategies for industrial air pollution sources using American Meteorological Society/Environmental Protection Agency Regulatory Model with simulated meteorology by Weather Research and Forecasting Model. J. Clean. Prod. 2015, 116, 110–117. [CrossRef]
- 43. Kumar, A.; Dikshit, A.K.; Fatima, S.; Patil, R.S. Application of WRF Model for Vehicular Pollution Modelling Using AERMOD. *Atmos. Clim. Sci.* 2015, *5*, 57–62. [CrossRef]
- Turtos Carbonell, L.; Mastrapa, G.C.; Rodriguez, Y.F.; Alvarez Escudero, L.; Sanchez Gacita, M.; Morlot, A.B.; Montejo, I.B.; Ruiz, E.M.; Rivas, S.P.; Carbonell, L.T. Assessment of the Weather Research and Forecasting model implementation in Cuba addressed to diagnostic air quality modeling. *Atmos. Pollut. Res.* 2013, *4*, 64–74. [CrossRef]
- 45. ul Haq, A.; Nadeem, Q.; Farooq, A.; Irfan, N.; Ahmad, M.; Ali, M.R. Assessment of AERMOD modeling system for application in complex terrain in Pakistan. *Atmos. Pollut. Res.* **2019**, *10*, 1492–1497. [CrossRef]
- Ruggeri, M.F.; Lana, N.B.; Altamirano, J.C.; Puliafito, S.E. Spatial distribution, patterns and source contributions of POPs in the atmosphere of Great Mendoza using the WRF/CALMET/CALPUFF modelling system. *Emerg. Contam.* 2020, *6*, 103–113. [CrossRef]
- Sunny Lim, K.S.; Lim, J.M.; Lee, J.; Shin, H.H. Impact of boundary layer simulation on predicting radioactive pollutant dispersion: A case study for HANARO research reactor using the WRF-MMIF-CALPUFF modeling system. *Nucl. Eng. Technol.* 2021, 53, 244–252. [CrossRef]
- Tartakovsky, D.; Stern, E.; Broday, D.M. Dispersion of TSP and PM10 emissions from quarries in complex terrain. *Sci. Total Environ.* 2016, 542, 946–954. [CrossRef] [PubMed]
- U.S. EPA. AERMOD Model Formulation and Evaluation; EPA-454/B-22-009, U.S. EPA, RTP, NC.; U.S. EPA: Washington, DC, USA, 2022.
- Giovannini, L.; Antonacci, G.; Zardi, D.; Laiti, L.; Panziera, L. Sensitivity of simulated wind speed to spatial resolution over complex terrain. *Energy Procedia* 2014, 59, 323–329. [CrossRef]

- 51. Nachamkin, J.E.; Cook, J.; Frost, M.; Martinez, D.; Sprung, G. Evaluation of dispersion forecasts driven by atmospheric model output at coarse and fine resolution. *J. Appl. Meteorol. Climatol.* **2007**, *46*, 1967–1980. [CrossRef]
- 52. U.S. EPA Support Center for Regulatory Atmospheric Modeling—Preferred and Recommended Models. Available online: https://www.epa.gov/scram/air-quality-dispersion-modeling-preferred-and-recommended-models (accessed on 11 September 2022).
- 53. Rzeszutek, M. Parameterization and evaluation of the CALMET/CALPUFF model system in near-field and complex terrain—Terrain data, grid resolution and terrain adjustment method. *Sci. Total Environ.* **2019**, *689*, 31–46. [CrossRef]
- Dresser, A.L.; Huizer, R.D. CALPUFF and AERMOD model validation study in the near field: Martins Creek revisited. J. Air Waste Manag. Assoc. 2011, 61, 647–659. [CrossRef] [PubMed]
- 55. Venkatram, A.; Brode, R.W.; Cimorelli, A.J.; Lee, R.F.; Paine, R.J.; Perry, S.G.; Peters, W.D.; Weil, J.C.; Wilson, R.B. A complex terrain dispersion model for regulatory applications. *Atmos. Environ.* **2001**, *35*, 4211–4221. [CrossRef]
- Perry, S.G.; Cimorelli, A.J.; Paine, R.J.; Brode, R.W.; Weil, J.C.; Venkatram, A.; Wilson, R.B.; Lee, R.F.; Peters, W.D. AERMOD: A Dispersion Model for Industrial Source Applications. Part II: Model Performance against 17 Field Study Databases. J. Appl. Meteorol. 2005, 44, 694–708. [CrossRef]
- NJDEP. Validation of CALPUFF in the Near-Field; EPA-HQ-OAR-2011-0081-0013; Bureau of Technical Services, Division of Air Quality, New Jersey Department of Environmental Protection: West Berlin, NJ, USA, 2010.
- U.S. EPA. AERMOD: Latest Features and Evaluation Results; EPA-454/B-03-001, U.S. EPA, RTP, NC.; U.S. EPA: Washington, DC, USA, 2003.
- Paumier, J.O.; Burns, D.J.; Perry, S.G. CTDMPLUS: A Dispersion Model for Sources near Complex Topography. Part II: Performance Characteristics. J. Appl. Meteorol. 1992, 31, 646–660. [CrossRef]
- 60. Banta, R.M.; Berri, G.; Blumen, W.; Carruthers, D.J.; Dalu, G.A.; Durran, D.R.; Egger, J.; Garratt, J.R.; Hanna, S.R.; Hunt, J.C.R.; et al. *Atmospheric Processes over Complex Terrain*; Blumen, W., Ed.; American Meteorological Society: Boston, MA, USA, 1990; Volume 23.
- 61. Truppi, L.E. EPA Complex Terrain Model Development: Description of a Computer Data Base from the Full Scale Plume Study, Tracy Power Plant, Nevada; EPA/600/S3-86/068, U.S. EPA, RTP, NC.; U.S. EPA: Washington, DC, USA, 1987.
- 62. Mesinger, F.; DiMego, G.; Kalnay, E.; Mitchell, K.; Shafran, P.C.; Ebisuzaki, W.; Jović, D.; Woollen, J.; Rogers, E.; Berbery, E.H.; et al. North American regional reanalysis. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 343–360. [CrossRef]
- Aylas, Y.G.R.; de Souza Campos Correa, W.; Santiago, A.M.; Reis Junior, N.C.; de Almeida Albuquerque, T.T.; Santos, J.M.; Moreira, D.M. Influence of land use on the performance of the WRF model in a humid tropical climate. *Theor. Appl. Climatol.* 2020, 141, 201–214. [CrossRef]
- 64. Jiménez-Esteve, B.; Udina, M.; Soler, M.R.; Pepin, N.; Miró, J.R. Land use and topography influence in a complex terrain area: A high resolution mesoscale modelling study over the Eastern Pyrenees using the WRF model. *Atmos. Res.* **2018**, 202, 49–62. [CrossRef]
- De Meij, A.; Vinuesa, J.F. Impact of SRTM and Corine Land Cover data on meteorological parameters using WRF. *Atmos. Res.* 2014, 143, 351–370. [CrossRef]
- 66. De Meij, A.; Bossioli, E.; Penard, C.; Vinuesa, J.F.; Price, I. The effect of SRTM and Corine Land Cover data on calculated gas and PM10 concentrations in WRF-Chem. *Atmos. Environ.* **2015**, *101*, 177–193. [CrossRef]
- 67. Mohan, M.; Sati, A.P. WRF model performance analysis for a suite of simulation design. Atmos. Res. 2016, 169, 280–291. [CrossRef]
- Jiménez, P.A.; Dudhia, J. Improving the Representation of Resolved and Unresolved Topographic Effects on Surface Wind in the WRF Model. J. Appl. Meteorol. Climatol. 2012, 51, 300–316. [CrossRef]
- Liu, Y.; Chen, X.; Li, Q.; Yang, J.; Li, L.; Wang, T. Impact of different microphysics and cumulus parameterizations in WRF for heavy rainfall simulations in the central segment of the Tianshan Mountains, China. *Atmos. Res.* 2020, 244, 105052. [CrossRef]
- Hong, S.Y.; Noh, Y.; Dudhia, J. A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Weather. Rev.* 2006, 134, 2318–2341. [CrossRef]
- Banks, R.F.; Tiana-Alsina, J.; Baldasano, J.M.; Rocadenbosch, F.; Papayannis, A.; Solomos, S.; Tzanis, C.G. Sensitivity of boundarylayer variables to PBL schemes in the WRF model based on surface meteorological observations, lidar, and radiosondes during the HygrA-CD campaign. *Atmos. Res.* 2016, 176–177, 185–201. [CrossRef]
- 72. Hong, S.; Lim, J. The WRF single-moment 6-class microphysics scheme (WSM6). J. Korean Meteorol. Soc. 2006, 42, 129–151.
- 73. Kain, J.S. The Kain—Fritsch convective parameterization: An update. J. Appl. Meteorol. 2004, 43, 170–181. [CrossRef]
- Chen, F.; Dudhia, J. Coupling an Advanced Land Surface Hydrology Model with the Penn State-NCAR MM5 Modeling System. Mon. Weather. Rev. 2001, 129, 17.
- Angevine, W.M.; Eddington, L.; Durkee, K.; Fairall, C.; Bianco, L.; Brioude, J. Meteorological model evaluation for CalNex 2010. Mon. Weather. Rev. 2012, 140, 3885–3906. [CrossRef]
- Borge, R.; Alexandrov, V.; José del Vas, J.; Lumbreras, J.; Rodríguez, E. A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula. *Atmos. Environ.* 2008, 42, 8560–8574. [CrossRef]
- 77. Mlawer, E.J.; Taubman, S.J.; Brown, P.D.; Iacono, M.J.; Clough, S.A. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. D Atmos.* **1997**, *102*, 16663–16682. [CrossRef]
- Dudhia, J. Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. J. Atmos. Sci. 1989, 46, 3077–3107. [CrossRef]

- 79. Karamchandani, P.; Emery, C.; Brashers, B. User's Manual: The Mesoscale Model Interface (MMIF) Program Version 4.0; U.S. EPA, RTP, NC.; U.S. EPA: Washington, DC, USA, 2022.
- Jiménez, P.A.; Dudhia, J. On the ability of the WRF model to reproduce the surface wind direction over complex terrain. J. Appl. Meteorol. Climatol. 2013, 52, 1610–1617. [CrossRef]
- Emery, C.; Tai, E.; Yarwood, G. Enhanced Meteorological Modeling and Performance Evaluation for Two Texas Ozone Episodes; Final Report Submitted to Texas Natural Resources Conservation Commission, prepared by ENVIRON.; International Corp.: Novato, CA, USA, 2001.
- 82. Carslaw, D.C.; Ropkins, K. Openair—An R package for air quality data analysis. *Environ. Model. Softw.* 2012, 27–28, 52–61. [CrossRef]
- 83. Willmott, C.J.; Robeson, S.M.; Matsuura, K. A refined index of model performance. Int. J. Climatol. 2012, 32, 2088–2094. [CrossRef]
- Kryza, M.; Wałaszek, K.; Ojrzyńska, H.; Szymanowski, M.; Werner, M.; Dore, A.J. High-Resolution Dynamical Downscaling of ERA-Interim Using the WRF Regional Climate Model for the Area of Poland. Part 1: Model Configuration and Statistical Evaluation for the 1981–2010 Period. *Pure Appl. Geophys.* 2017, 174, 511–526. [CrossRef]
- Ojrzyńska, H.; Kryza, M.; Wałaszek, K.; Szymanowski, M.; Werner, M.; Dore, A.J. High-Resolution Dynamical Downscaling of ERA-Interim Using the WRF Regional Climate Model for the Area of Poland. Part 2: Model Performance with Respect to Automatically Derived Circulation Types. *Pure Appl. Geophys.* 2017, 174, 527–550. [CrossRef]
- Zhang, H.; Chen, G.; Hu, J.; Chen, S.H.; Wiedinmyer, C.; Kleeman, M.; Ying, Q. Evaluation of a seven-year air quality simulation using the Weather Research and Forecasting (WRF)/Community Multiscale Air Quality (CMAQ) models in the eastern United States. *Sci. Total Environ.* 2014, 473–474, 275–285. [CrossRef] [PubMed]
- Wang, L.; Zhang, Y.; Wang, K.; Zheng, B.; Zhang, Q.; Wei, W. Application of Weather Research and Forecasting Model with Chemistry (WRF/Chem) over northern China: Sensitivity study, comparative evaluation, and policy implications. *Atmos. Environ.* 2014, 124, 337–350. [CrossRef]
- 88. U.S. EPA. Modeling Guidance for Demonstrating Air Quality Goals for Ozone, PM2.5, and Regional Haze; EPA 454/R-18-009, U.S. EPA, RTP, NC.; U.S. EPA: Washington, DC, USA, 2018.
- 89. Canty, A.; Ripley, B.D. *Boot: Bootstrap R (S-Plus) Functions, R Package Version 1.3-28.1;* R Foundation for Statistical Computing: Vienna, Austria, 2022.
- 90. Davison, A.C.; Hinkley, D.V. Bootstrap Methods and Their Applications; Cambridge University Press: Cambridge, UK, 1997.
- 91. R Core Team. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2020.
- 92. R Studio Team. RStudio: Integrated Development Environment for R.; R Studio Team: Boston, MA, USA, 2020.
- 93. Chang, J.C.; Hanna, S.R. *Technical Descriptions and User's Guide for the BOOT Statistical Model Evaluation Software Package, Version* 2.0; George Mason University and Harvard School of Public Health: Fairfax, VA, USA, 2005.
- U.S. EPA. Protocol for Determining the Best Performing Model; EPA-454/R-92-025, U.S. EPA, RTP, NC.; U.S. EPA: Washington, DC, USA, 1992.
- 95. Lingard, J.; Labrador, L.; Brookes, D.; Fraser, A. Statistical Evaluation of the Input Meteorological Data Used for the UK Air Quality Forecast (UK-AQF): RMP/1902: Defra Air Quality Forecasting; RICARDO-AEA/R/3388 Issue 1; DEFRA: London, UK, 2013.
- 96. Lai, Z.; Ma, R.; Huang, M.; Chen, C.; Chen, Y.; Xie, C.; Beardsley, R.C. Downwelling wind, tides, and estuarine plume dynamics. J. *Geophys. Res. Ocean.* 2016, 121, 4245–4263. [CrossRef]
- 97. Chen, F.; Miao, S.; Tewari, M.; Bao, J.W.; Kusaka, H. A numerical study of interactions between surface forcing and sea breeze circulations and their effects on stagnation in the greater Houston area. *J. Geophys. Res. Atmos.* **2011**, *116*, D12105. [CrossRef]
- 98. Matsuzaki, Y.; Fujiki, T.; Kawaguchi, K.; Inoue, T.; Iwamoto, T. Application of the WRF model to the coastal area at Ise Bay, Japan: Evaluation of model output sensitivity to input data. *Coast. Eng. J.* **2021**, *63*, 17–31. [CrossRef]
- Lucas-Picher, P.; Laprise, R.; Winger, K. Evidence of added value in North American regional climate model hindcast simulations using ever-increasing horizontal resolutions. *Clim. Dyn.* 2017, 48, 2611–2633. [CrossRef]
- 100. El-Samra, R.; Bou-Zeid, E.; El-Fadel, M. To what extent does high-resolution dynamical downscaling improve the representation of climatic extremes over an orographically complex terrain? *Theor. Appl. Climatol.* **2018**, *134*, 265–282. [CrossRef]
- 101. Zhang, H.; Pu, Z.; Zhang, X. Examination of errors in near-surface temperature and wind from WRF numerical simulations in regions of complex terrain. *Weather. Forecast.* **2013**, *28*, 893–914. [CrossRef]
- Szulecka, A.; Bogacki, M. Evaluation of available data initializing retrospective simulations using WRF meteorological model. Proc. ECOpole 2018, 12, 77–86.
- Madala, S.; Satyanarayana, A.N.V.; Srinivas, C.V.; Kumar, M. Mesoscale atmospheric flow-field simulations for air quality modeling over complex terrain region of Ranchi in eastern India using WRF. *Atmos. Environ.* 2015, 107, 315–328. [CrossRef]
- Hanna, S.R.; Yang, R. Evaluations of Mesoscale Models' Simulations of Near-Surface Winds, Temperature Gradients, and Mixing Depths. J. Appl. Meteorol. Climatol. 2001, 40, 1095–1104. [CrossRef]
- Briggs, G.A. Plume Rise Predictions. In *Lectures on Air Pollution and Environmental Impact Analyses*; American Meteorological Society: Boston, MA, USA, 1982; pp. 59–111.
- 106. Liu, Y.; Warner, T.T.; Astling, E.G.; Bowers, J.F.; Davis, C.A.; Halvorson, S.F.; Rife, D.L.; Sheu, R.S.; Swerdlin, S.P.; Xu, M. The operational mesogamma-scale analysis and forecast system of the U.S. Army Test and Evaluation Command. Part II: Interrange comparison of the accuracy of model analyses and forecasts. *J. Appl. Meteorol. Climatol.* 2008, 47, 1093–1104. [CrossRef]

- 107. Witha, B.; Hahmann, A.; Sile, T.; Dörenkämper, M.; Ezber, Y.; García-Bustamante, E.; González-Rouco, J.F.; Leroy, G.; Navarro, J. WRF Model Sensitivity Studies and Specifications for the NEWA Mesoscale Wind Atlas Production Runs: Deliverable D4.3; NEWA—New European Wind Atlas; DTU: Copenhagen, Denmark, 2019.
- 108. Wiersema, D.J.; Lundquist, K.A.; Chow, F.K. A Framework for WRF to WRF-IBM Grid Nesting to Enable Multiscale Simulations; LLNL-SR-704084; Lawrence Livermore National Lab: Livermore, CA, USA, 2016.
- 109. Cox, W.M.; Tikvart, J.A. A statistical procedure for determining the best performing air quality simulation model. *Atmos. Environment. Part A General Top.* **1990**, *24*, 2387–2395. [CrossRef]
- 110. Hanna, S.; Chang, J. Acceptance criteria for urban dispersion model evaluation. *Meteorol. Atmos. Phys.* **2012**, *116*, 133–146. [CrossRef]
- Van Ulden, A.P.; Holtslag, A.A.M. Estimation of Atmospheric Boundary Layer Parameters for Diffusion Applications. J. Appl. Meteorol. Climatol. 1985, 24, 1196–1207. [CrossRef]
- 112. Businger, J.A.; Wyngaard, J.C.; Izumi, Y.; Bradley, E.F. Flux-Profile Relationships in the Atmospheric Surface Layer. *J. Atmos. Sci.* **1971**, *28*, 181–189. [CrossRef]
- 113. Rai, R.K.; Berg, L.K.; Kosović, B.; Haupt, S.E.; Mirocha, J.D.; Ennis, B.L.; Draxl, C. Evaluation of the impact of horizontal grid spacing in terra incognita on coupled mesoscale-microscale simulations using the WRF framework. *Mon. Weather. Rev.* 2019, 147, 1007–1027. [CrossRef]
- Emmanouil, G.; Vlachogiannis, D.; Sfetsos, A. Exploring the ability of the WRF-ARW atmospheric model to simulate different meteorological conditions in Greece. *Atmos. Res.* 2021, 247, 105226. [CrossRef]
- 115. Carvalho, D.; Rocha, A.; Gómez-Gesteira, M.; Santos, C. A sensitivity study of the WRF model in wind simulation for an area of high wind energy. *Environ. Model. Softw.* **2012**, *33*, 23–34. [CrossRef]
- 116. Gopalan, H.; Gundling, C.; Brown, K.; Roget, B.; Sitaraman, J.; Mirocha, J.D.; Miller, W.O. A coupled mesoscale-microscale framework for wind resource estimation and farm aerodynamics. *J. Wind Eng. Ind. Aerodyn.* **2014**, *132*, 13–26. [CrossRef]
- 117. Carvalho, D.; Rocha, A.; Santos, C.S.; Pereira, R. Wind resource modelling in complex terrain using different mesoscale-microscale coupling techniques. *Appl. Energy* **2013**, *108*, 493–504. [CrossRef]
- 118. Kadaverugu, R.; Purohit, V.; Matli, C.; Biniwale, R. Improving accuracy in simulation of urban wind flows by dynamic downscaling WRF with OpenFOAM. *Urban Clim.* **2021**, *38*, 100912. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.