https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


# A terrain-stratified comparison of regional MPAS-A (v8.3.1) and nested WRF (v4.7.1) for short-term near-surface wind forecasting

Ruhi Deniz Yalcin[1], Mustafa Tugrul Yilmaz[1], Ismail Yucel[1], Omer Lutfi Sen[2], Sertac Oruc[3], Ali Ulvi Galip Senocak[3]

[1]Department of Civil Engineering, Middle East Technical University, Ankara, Türkiye
[2]Eurasia Institute of Earth Sciences, Istanbul Technical University, Istanbul, Türkiye
[3]Department of Civil Engineering, Ankara Yıldırım Beyazıt University, Ankara, Türkiye

*Correspondence to*: Ruhi Deniz Yalcin (deniz.yalcin@metu.edu.tr)

**Abstract.** Accurate deterministic forecasts of 10 m wind over complex terrain remain challenging, and the relative performance of nested regional and variable-resolution modelling strategies is still insufficiently documented. This study presents a controlled comparison of five deterministic forecast products over Türkiye: four nested WRF configurations at 12 km parent and 4 km child domain resolution with one-way and two-way nesting, and regional MPAS-Atmosphere with an approximately 4 km refined inner region. All experiments use identical 0.25° GFS forcing, are initialised daily at 00:00 UTC, and integrated for 48 h. After excluding the first 6 h as spin-up, verification is performed over lead times 6–47 h using 10 m wind speed observations from 509 TSMS stations during January and June 2023. Terrain effects are explicitly addressed through a Terrain Ruggedness Index computed from SRTM GL3 elevation data and a three-class terrain stratification. Model performance is assessed using mean error, mean absolute error, root mean square error, and Pearson correlation coefficient.

MPAS achieves the lowest error across all aggregate metrics, with ME of 0.76 m s−1 and RMSE of 2.16 m s−1, compared to 0.99 m s−1 and 2.27 m s−1 for the best WRF configuration. All models exhibit systematic positive bias. The performance gap between MPAS and WRF is strongly terrain-dependent: negligible in low-complexity terrain but reaching 41 % ME reduction at high-complexity sites. Increasing WRF resolution from 12 km to 4 km does not consistently reduce bias, with the coarser 12 km two-way configuration outperforming the finer 4 km two-way configuration across both months and all terrain classes, pointing to grey-zone boundary layer limitations and lateral boundary constraints. Two-way nesting provides meaningful improvement at 12 km but offers diminishing returns at 4 km. MPAS's aggregate ME advantage partly reflects a compensating bias structure: reduced overprediction at calm and moderate wind speeds is offset by stronger underprediction at high wind speeds, reaching ME of −3.28 m s−1 at high-complexity sites. This wind speed dependence also explains the spatial reversal along the Aegean coast, where WRF 4 km configuration outperforms MPAS at 70 % of stations. These results indicate that neither modelling approach offers a universal advantage; relative performance depends on terrain complexity, local wind climate, and the verification metric of interest.

**Keywords:** WRF, MPAS-A, wind speed, complex terrain, TRI, deterministic forecast verification

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


## 1 Introduction

Accurate short-term forecasting of near-surface (10 m) wind is essential for a wide range of weather-sensitive decisions, including wind-power integration, grid balancing, reserve allocation, aviation, maritime operations, and broader operational risk management. Day-ahead and intraday forecast horizons are particularly important because they directly affect dispatch planning and the management of imbalance costs in power systems (Costa et al., 2008; Foley et al., 2012; Hanifi et al., 2020). Despite sustained advances in numerical weather prediction (NWP), achieving consistently high skill for 10 m wind forecasts remains difficult in regions where the flow is strongly modulated by topography (Jiménez and Dudhia, 2012; Gómez-Navarro et al., 2015; Chen et al., 2020).

Complex terrain is a setting in which multiple atmospheric processes interact across scales, many of which are difficult to resolve explicitly and equally difficult to represent through parameterisation. Mountain environments can host valley circulations, gap flows, channeling, slope-wind systems, and sharp local contrasts in boundary-layer structure, all of which influence near-surface wind variability and predictability (Whiteman, 2000). Even at kilometre-scale grid spacing, unresolved orographic detail and physics choices, especially in the planetary boundary layer and surface-layer formulations, can introduce systematic wind-speed errors (Jiménez and Dudhia, 2012; Gómez-Navarro et al., 2015; Chen et al., 2020). These limitations are particularly relevant for wind-energy applications, where terrain-driven forecast errors can propagate directly into resource assessment and operational decision-making (Barber et al., 2022).

A common operational strategy for representing regional detail is to employ the Weather Research and Forecasting (WRF) model with nested domains, in which a coarse outer grid supplies the large-scale flow and one or more inner grids increase resolution over the area of interest (Skamarock et al., 2019). Although this approach is computationally efficient and widely used, it also has well-documented limitations. Regional solutions remain sensitive to the lateral boundary conditions imposed by the driving model, and inconsistencies at the nest interface may alter the evolution of mesoscale structures and error growth (Davies, 1976; Warner et al., 1997; Harris and Durran, 2010). These issues are especially important for near-surface winds over heterogeneous terrain, where small-scale topographic controls can dominate local variability.

Variable-resolution modelling offers an alternative strategy that avoids conventional nested lateral boundaries within the region of interest. The Model for Prediction Across Scales – Atmosphere (MPAS-A) employs an unstructured centroidal Voronoi mesh with C-grid staggering, allowing a high-resolution region to be embedded within a single global mesh with smooth transitions in horizontal resolution (Skamarock et al., 2012). Previous idealised and method-oriented studies have shown that the numerical treatment of mesh refinement, and the strategies used to minimise refinement-related artefacts, are central to the fidelity of variable-resolution simulations (Park et al., 2014; Du and Stechmann, 2023). Comparative studies have also examined MPAS-A and WRF in selected configurations and case studies, indicating that their relative performance depends on flow regime, terrain, and configuration choices (Hagos et al., 2013; Kramer et al., 2020; Gürer et al., 2024; Shin et al., 2025). Most recently, Shin et al. (2025) compared regional MPAS and WRF over the Korean Peninsula and found broadly similar performance between the two models across complex terrain, without a systematic terrain-dependent divergence.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


Whether such divergence emerges over a geographically more diverse domain, and how it relates to terrain complexity, remains an open question that the present study addresses.

Despite this growing literature, systematic head-to-head comparisons between variable-resolution MPAS-A and nested WRF remain limited for deterministic short-range near-surface wind forecasting over geographically diverse complex terrain. This gap is especially relevant for Türkiye, where coastal plains, enclosed basins, plateaus, and major mountain belts coexist within relatively short distances and where terrain-controlled wind regimes are expected to vary sharply across space (Şahin and Türkeş, 2013; Çetin et al., 2024). A comparison over Türkiye therefore provides a demanding test bed for evaluating how model design, refinement strategy, and terrain complexity interact in short-range wind prediction.

In response to this gap, the present study compares five deterministic forecast products over Türkiye: WRF d01 one-way (12 km), WRF d01 two-way (12 km), WRF d02 one-way (4 km), WRF d02 two-way (4 km), and MPAS-A with an approximately 4 km refined inner region over Türkiye, where d01 and d02 denote the WRF outer (12 km) and inner (4 km) domains respectively, as used throughout this study. All experiments are driven by the same 0.25° Global Forecast System (GFS) analyses and forecasts (National Centers for Environmental Prediction et al., 2015), are initialised daily at 00:00 UTC, and are integrated for 48 h. After excluding the first 6 h as spin-up, verification is performed over lead times 6–47 h. The evaluation is based on 10 m wind speed observations from 509 Turkish State Meteorological Service (TSMS) stations during January and June 2023.

To avoid masking terrain-dependent behaviour through broad spatial averaging, the verification is stratified by three terrain-complexity classes derived from station-representative Terrain Ruggedness Index (TRI) values computed from the SRTM GL3 digital elevation model (Riley et al., 1999; Farr et al., 2007 ; NASA JPL, 2013). Forecast quality is assessed using four deterministic metrics: mean error (ME), mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient (CC). This design links the model intercomparison directly to terrain controls on forecast skill and enables a consistent assessment of how nesting strategy, variable-resolution mesh design, and topographic complexity shape short-range 10 m wind performance.

The contribution of this study lies in its controlled and explicitly terrain-aware comparison design. It evaluates matched WRF and MPAS-A experiments under common large-scale forcing, a shared physics suite, identical verification periods, and a dense station network, while retaining one-way and two-way nesting as explicit components of the WRF benchmark. This makes it possible to identify where differences are associated with the horizontal discretisation strategy itself and where they are amplified or reduced by terrain complexity. The remainder of the paper is organised as follows. Section 2 describes the observational dataset, model configurations, terrain-complexity classification, and verification framework. Section 3 presents the results, Sect. 4 discusses the underlying mechanisms, and Sect. 5 summarises the main conclusions.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


## 2 Methodology

### 2.1 Observational data

The observational dataset consists of 10 m wind speed observations from the Turkish State Meteorological Service (TSMS) surface-station network. Two contrasting months were selected for the evaluation period: January 2023, representing winter synoptic conditions, and June 2023, representing warm-season conditions dominated by the interplay between regional synoptic forcing and localised thermally driven circulation. From the full TSMS archive of 1669 station identifiers, stations were first screened using a physical-range filter retaining only wind-speed values within the 0–50 m s$^{-1}$ interval. Stations with insufficient temporal coverage were subsequently excluded by applying a minimum data availability threshold of 95 %, defined as the fraction of expected hourly observation counts across both evaluation months combined. This two-step procedure yielded a final set of 509 stations.

Observational quality control was limited to the physical-range filter described above, with no automated spike removal, spatial consistency checks, or homogenisation procedures applied. This is consistent with common practice in kilometre-scale NWP verification studies where the focus is on aggregate model performance across a dense station network rather than on the quality of individual time series (Jiménez and Dudhia, 2012). Each verified daily cycle yields up to 42 matched hourly observation–forecast pairs, corresponding to a maximum of 1302 pairs per station in January (31 days) and 1260 in June (30 days).

### 2.2 Model setup

The modelling framework compares a nested limited-area WRF configuration with a variable-resolution regional MPAS-Atmosphere (MPAS-A) configuration for short-range near-surface wind forecasting over Türkiye. Both models were driven by 0.25° Global Forecast System (GFS) analyses and forecasts (National Centers for Environmental Prediction et al., 2015). Simulations were initialised daily at 00:00 UTC, with lateral boundary conditions updated every 3 h for both WRF and MPAS. Each daily cycle consisted of a 48 h integration; the first 6 h were discarded as model spin-up, a common practice in short-range NWP verification to allow the model dynamics and physics to adjust to the initial conditions before evaluation begins. This choice is further supported by the practical consideration that GFS output for the 00:00 UTC cycle becomes available around 03:30 UTC, meaning that real-time forecasts would in any case only be issued from around that time onward. Verification was therefore performed over lead times 6–47 h, yielding 42 hourly evaluation windows per cycle. Model output was archived at 10 min intervals in NetCDF format and subsequently aggregated to hourly means for comparison with observations. The subsequent evaluation is restricted to deterministic verification.

The regional WRF simulations were performed with WRF-ARW, described in Skamarock et al. (2019); here version 4.7.1 was used. The MPAS simulations used the variable-resolution solver described in Skamarock et al. (2012); here version 8.3.1 was used. Both models employ a shared physics suite summarised in Table 1. The Grell–Freitas scale-aware convective parameterisation was kept active on all domains, including the 4 km grids, following its design as a scale-aware scheme that

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


progressively reduces parameterised convective transport as the grid spacing decreases (Grell and Freitas, 2014). Model configurations are summarised in Table 2.

Both models use 55 vertical levels extending from the surface to a model top of 50 hPa for WRF and 21 km for MPAS. WRF employs a hybrid terrain-following (η) vertical coordinate, while MPAS uses height-based vertical levels with smoothed terrain-following coordinate surfaces (Klemp, 2011).

**Table 1: Shared physics parameterisation suite applied to all WRF and MPAS configurations in this study.**

| Parameterisation | Scheme | Reference |
|---|---|---|
| Microphysics | Thompson aerosol-aware | Thompson and Eidhammer (2014) |
| Convection | Grell–Freitas (all domains) | Grell and Freitas (2014) |
| PBL / surface layer | MYNN level 2.5 | Nakanishi and Niino (2009) |
| Radiation (SW and LW) | RRTMG | Iacono et al. (2008) |
| Land surface | Noah | Tewari et al. (2004) |

**Table 2: Summary of model configuration parameters for the WRF nested domains and MPAS variable-resolution mesh used in this study.**

| Parameter | WRF d01 | WRF d02 | MPAS inner | MPAS outer |
|---|---|---|---|---|
| Horizontal resolution | 12 km | 4 km | ~4 km | ~12 km |
| Time step | 60 s | 20 s | 20 s | 20 s |
| Vertical levels | 55 | 55 | 55 | 55 |
| Model top | 50 hPa | 50 hPa | 21 km | 21 km |
| Vertical coordinate | Hybrid η | Hybrid η | Height-based | Height-based |
| Nesting | Parent | 1-way / 2-way | Single mesh | Single mesh |
| Software version | WRF v4.7.1 | WRF v4.7.1 | MPAS v8.3.1 | MPAS v8.3.1 |

### 2.3 Nested WRF model

The regional simulations were performed with the Advanced Research WRF (ARW) core, which solves the fully compressible, nonhydrostatic Euler equations on an Arakawa C grid with a hybrid terrain-following vertical coordinate (Skamarock et al., 2019). The configuration uses a two-domain nesting setup centred over central Türkiye (39° N, 34° E, Lambert conformal projection): a 12 km parent domain (d01) covering 300 × 200 grid points (approximately 3588 km × 2388 km) and a 4 km nested domain (d02) of 481 × 289 grid points (approximately 1920 km × 1152 km) spanning Türkiye and its surrounding region (Fig. 1). The nested domain originates at grid point (72, 54) of the parent domain, providing a buffer of at least 50 grid

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


cells on all sides between the inner domain boundary and the outer domain edge, consistent with the guidance of Davies (1976) and Warner et al. (1997) for minimising lateral boundary artefacts in the nested domain solution.

The experimental design retains both one-way and two-way nesting configurations, yielding four distinct WRF products: d01 one-way (12 km), d01 two-way (12 km), d02 one-way (4 km), and d02 two-way (4 km). This design keeps the nesting strategy as an explicit element of the intercomparison, since one-way versus two-way feedback can alter error propagation and the representation of mesoscale structures near the nest boundary (Harris and Durran, 2010).


**Figure 1: WRF model domain configuration. The outer blue boundary denotes the 12 km parent domain (d01); the inner red boundary denotes the 4 km nested domain (d02).**

### 2.4 Regional MPAS model

The MPAS experiments were carried out on a variable-resolution global mesh constructed from centroidal Voronoi tessellations with C-grid staggering (Skamarock et al., 2012). The mesh was generated using the JIGSAW mesh generator (Engwirda, 2017) interfaced through the jigsawpy Python bindings and the MPAS-Tools library.

The target resolution field was defined on a global 0.1° latitude–longitude grid and mapped to a normalised elliptical distance coordinate centred at 39° N, 35° E. The ellipse semi-axes were set to $R_a$ = 850 km in the zonal direction and $R_b$ = 600 km in the meridional direction, reflecting the elongated east–west extent of Türkiye. The normalised distance from the ellipse centre was computed as follows:

$$d = \sqrt{\left(\frac{x_{hs}}{R_a}\right)^2 + \left(\frac{y_{hs}}{R_b}\right)^2} \quad (1)$$

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


where $x_{hs}$ and $y_{hs}$ are the great-circle distances in the zonal and meridional directions respectively, derived via the haversine formula. A piecewise linear interpolation function then mapped this normalised distance to a target resolution: 4 km for $d \leq 1.0$ (the inner quasi-uniform core), a smooth linear transition from 4 to 12 km for $1.0 < d \leq 1.54$, and 12 km for $d > 1.54$ out to the clipping boundary at $d = 2.2$. The smooth, gradual transition between resolution zones is a defining feature of the MPAS variable-resolution approach and avoids the sharp internal-resolution jumps inherent in conventional static nesting (Park et al., 2014).

The JIGSAW optimisation was run with a cell quality limit of 0.9375 and up to 100 optimisation iterations, producing a high-quality Voronoi tessellation throughout the transition region (Engwirda, 2017). The resulting triangulated mesh was converted to MPAS NetCDF format via jigsaw_to_netcdf and the MPAS-Tools convert utility. To configure MPAS as a regional model analogous to the WRF nested domain setup, the global mesh was subsequently clipped to a domain bounded by the outer ellipse ($R_a \times 2.2 = 1870$ km, $R_b \times 2.2 = 1320$ km), as shown in Fig. 2, using the create_region tool from MPAS-Limited-Area (https://github.com/MPAS-Dev/MPAS-Limited-Area). The mesh generation script is archived in the Zenodo repository associated with this study.

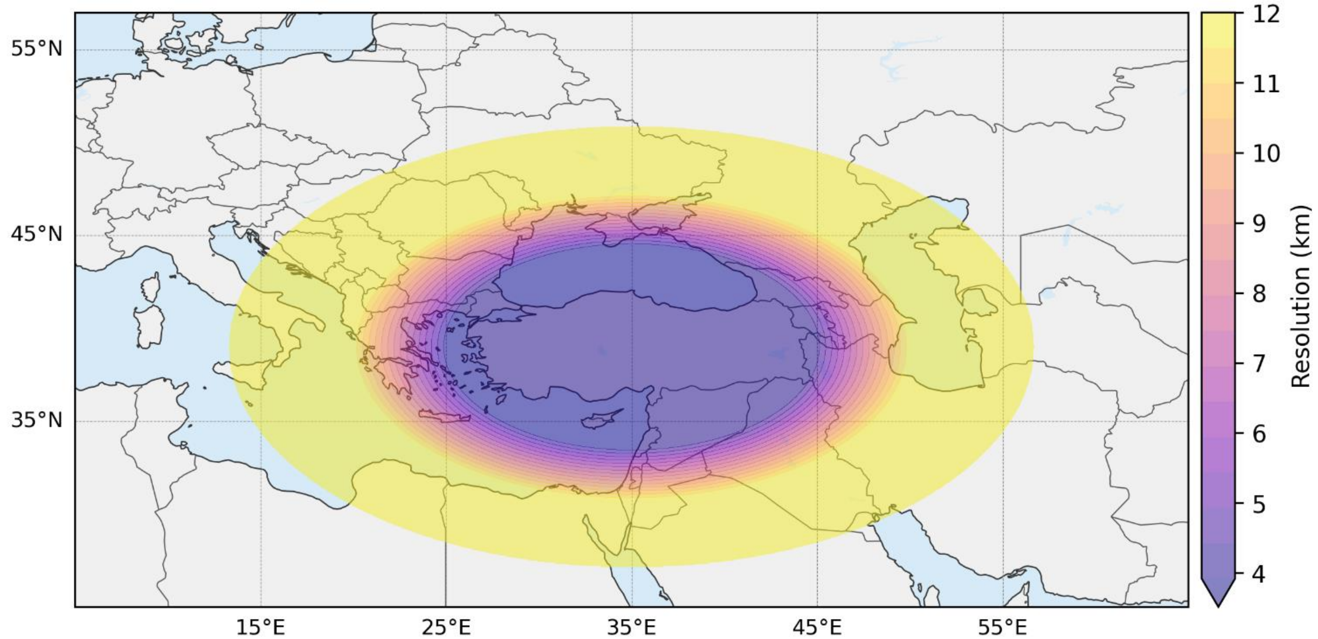


**Figure 2: MPAS variable-resolution mesh structure showing the quasi-uniform ~4 km inner core over Türkiye (dark region), the 4–12 km smooth transition zone, and the quasi-uniform ~12 km outer region.**

### 2.5 Terrain complexity classification

Türkiye exhibits pronounced topographic contrasts, from the low-lying coastal plains along the Aegean and Mediterranean to the rugged Northern Anatolian mountains along the Black Sea coast, the Taurus Range in the south, and the elevated Central Anatolian plateau in between (Fig. 3). Because terrain complexity is a well-known driver of near-surface wind forecast error

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


in kilometre-scale NWP (Jiménez and Dudhia, 2012; Gómez-Navarro et al., 2015), terrain ruggedness was introduced as an explicit stratification variable for the verification analysis.

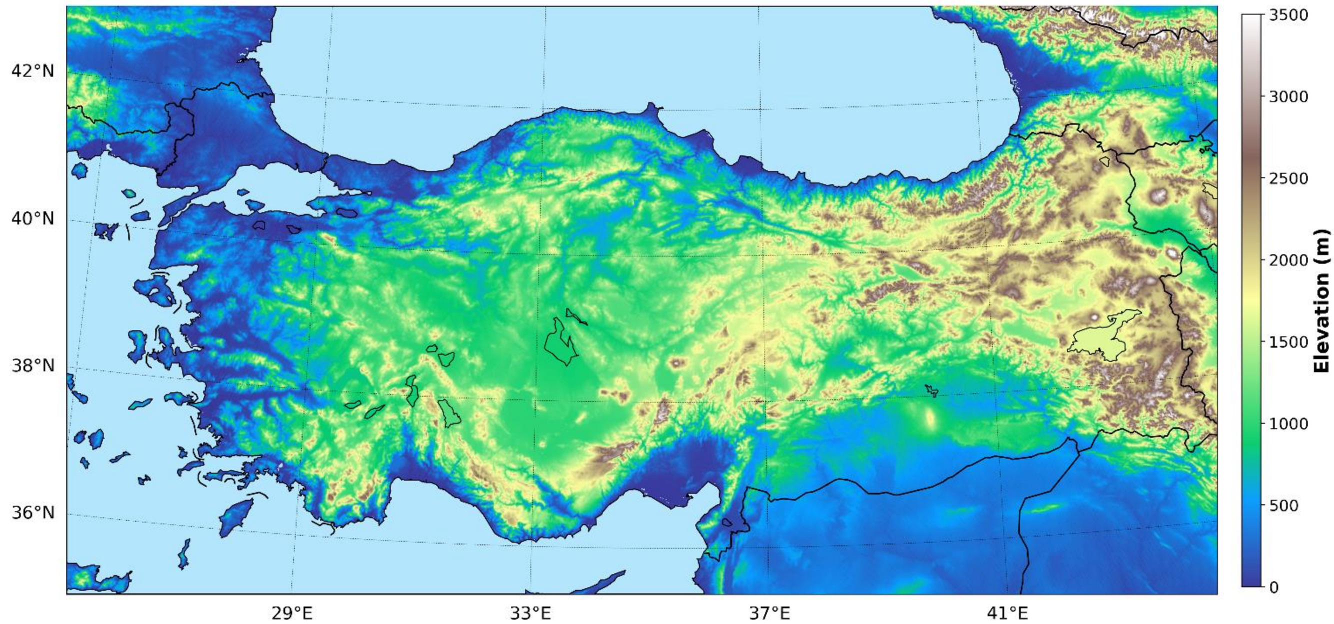


**Figure 3: Digital elevation model of the study domain derived from SRTM GL3 (90 m resolution). Elevation is shown in metres above sea level.**

### 2.5.1 Terrain Ruggedness Index (TRI) calculation

Terrain complexity was quantified using the Terrain Ruggedness Index (TRI) concept (Riley et al., 1999) computed from the Shuttle Radar Topography Mission (SRTM GL3) 90 m digital elevation model (Farr et al., 2007; NASA JPL, 2013). The TRI was computed for each DEM grid cell over a 10×10 pixel neighbourhood (approximately 900 m × 900 m at 90 m resolution) as the mean absolute elevation difference between the central cell and all surrounding cells within the window:

$$\mathrm{TRI}_{\mathrm{cell}} = \frac{1}{M} \sum_i |z_i - z_0|, \qquad (2)$$

where $z_0$ is the elevation of the central cell, $z_i$ is the elevation of each neighbouring cell within the window, and M is the number of valid cells in the neighbourhood. This formulation differs from the original 3×3 root-mean-square definition of Riley et al. (1999) in two respects: it uses mean absolute differences rather than root-mean-square differences, and it employs a larger 10×10 window to capture terrain heterogeneity at spatial scales more representative of the 4–12 km NWP grids evaluated in this study.

A station-representative TRI value was then assigned to each of the 509 evaluation stations by extracting the TRI value at the DEM grid cell nearest to the station location.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


### 2.5.2 Classification scheme

Three terrain-complexity classes were defined based on the distribution of station-representative TRI values over the study domain: low (TRI < 10), medium (10 ≤ TRI < 30), and high (TRI ≥ 30). The thresholds were selected to yield an approximately balanced distribution of stations across the three classes while reflecting natural breaks in the TRI distribution of the station network. These classes are used consistently throughout the stratified verification analysis.

### 2.5.3 Station distribution by complexity

Applying the classification thresholds to the 509-station evaluation set yields 159 low-complexity stations (31.2 %), 196 medium-complexity stations (38.5 %), and 154 high-complexity stations (30.3 %), as summarised in Table 3. The spatial distribution of TRI values across the study domain is shown in Fig. 4, where the highest ruggedness values correspond to the Northern Anatolian mountains along the Black Sea coast, the Taurus Range in southern Türkiye, and the volcanic highlands of eastern Anatolia.

**Table 3: Distribution of the 509 evaluation stations across terrain complexity classes defined by the station-representative Terrain Ruggedness Index (TRI).**

| Class | TRI range | Stations | Percentage |
|---|---|---|---|
| Low | < 10 | 159 | 31.2 % |
| Medium | 10–30 | 196 | 38.5 % |
| High | ≥ 30 | 154 | 30.3 % |
| Total | | 509 | 100 % |

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


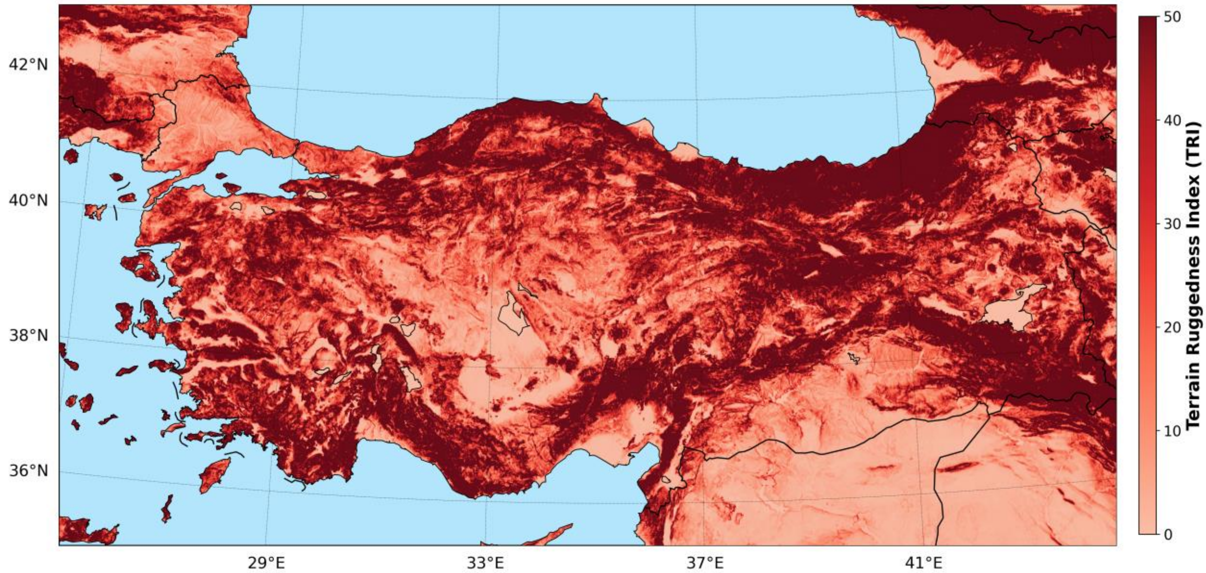


**Figure 4: Station-representative Terrain Ruggedness Index (TRI) over the study domain computed from SRTM GL3 90 m data using a 10×10 pixel (approximately 900 m × 900 m) neighbourhood, extracted at the nearest grid cell to each evaluation station.**

### 2.6 Verification metrics

Modelled 10 m wind speed was extracted from the grid-cell centre to each meteorological station, without horizontal interpolation, from 10 min output and then aggregated to hourly means prior to verification. Five deterministic forecast products are evaluated: WRF d01 one-way (hereafter WRF-d01-1way), WRF d01 two-way (hereafter WRF-d01-2way), WRF d02 one-way (hereafter WRF-d02-1way), WRF d02 two-way (hereafter WRF-d02-2way), and MPAS (~4 km over Türkiye). All verification was carried out using a custom Python processing pipeline.

The verification score set comprises four standard deterministic metrics (Murphy and Winkler, 1987; Jolliffe and Stephenson, 2012; Wilks, 2011). Mean error (ME) measures systematic bias:

$$ME = \frac{1}{N} \sum_{i=1}^{N} (F_i - O_i), \tag{3}$$

Mean absolute error (MAE) quantifies typical error magnitude:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |F_i - O_i|, \tag{4}$$

Root mean square error (RMSE) emphasises larger errors:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (F_i - O_i)^2}, \tag{5}$$

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


The Pearson correlation coefficient (CC) measures the linear temporal agreement between forecast and observed time series:

$$CC = \frac{\sum_{i=1}^{N}(F_i - \bar{F})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N}(F_i - \bar{F})^2}\sqrt{\sum_{i=1}^{N}(O_i - \bar{O})^2}}, \tag{6}$$

In the above, $F_i$ and $O_i$ denote forecast and observed hourly wind speeds, respectively; $\bar{F}$ and $\bar{O}$ denote the mean forecast and observation, respectively; and N is the number of matched pairs. All metrics are computed per station and per evaluation period (January, June, and the combined two-month set) and subsequently aggregated across the TRI-stratified subsets defined in Sect. 2.5. No probabilistic or ensemble-based scores were considered.

### 2.7 Computational environment

All simulations were executed on the TRUBA Hamsi computing cluster operated by TÜBİTAK ULAKBİM. WRF experiments used a single compute node (54 threads), while MPAS experiments used two nodes (108 threads). Model output was stored in NetCDF format at 10 min temporal resolution. In addition to 10 m wind speed, supplementary hydrometeorological variables (precipitation, temperature, humidity) were archived for potential future analysis. Total output volumes are approximately 108 GB for MPAS and 169 GB for the combined WRF domains (12 km and 4 km).

## 3 Results

### 3.1 Overall model performance

Table 4 presents the combined two-month verification statistics for all five model configurations across 509 stations. MPAS records the lowest ME (0.76 m s−1), MAE (1.56 m s−1), and RMSE (2.16 m s−1), and the highest CC (0.541) among all configurations. Correlation coefficients range narrowly from 0.47 to 0.54, which indicates that all models capture the temporal structure of wind variability with comparable skill regardless of discretisation strategy. More notable differences are observed in the error metrics.

All five configurations exhibit a positive bias, consistent with the systematic overprediction of near-surface wind speeds reported in kilometre-scale NWP studies over complex terrain (Jiménez and Dudhia, 2012; Gómez-Navarro et al., 2015; Chen et al., 2020). MPAS reduces this bias by 23 % relative to WRF-d02-2way (ME: 0.99 m s−1) and by 45 % relative to WRF-d01-1way (ME: 1.38 m s−1). Notably, increasing horizontal resolution from 12 km to 4 km does not reduce the positive bias: WRF-d02-2way (ME: 0.99 m s−1) performs worse than WRF-d01-2way (ME: 0.92 m s−1) in this regard, indicating that factors beyond resolution drive the systematic wind speed errors. The role of nesting strategy and terrain complexity in this behaviour is examined further in Sect. 3.2 and 3.5.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


**Table 4: Combined two-month verification statistics for all five model configurations across 509 TSMS stations. ME: mean error; MAE: mean absolute error; RMSE: root mean square error; CC: Pearson correlation coefficient. Statistics are aggregated over N = 1 293 039 matched observation–forecast pairs for lead times of 6–47 h.**

| | WRF-d01-1way | WRF-d01-2way | WRF-d02-1way | WRF-d02-2way | MPAS |
|---|---|---|---|---|---|
| CC | 0.471 | 0.530 | 0.532 | 0.536 | 0.541 |
| ME (m s−1) | 1.38 | 0.92 | 1.02 | 0.99 | 0.76 |
| MAE (m s−1) | 1.96 | 1.65 | 1.68 | 1.66 | 1.56 |
| RMSE (m s−1) | 2.60 | 2.25 | 2.30 | 2.27 | 2.16 |

## 3.2 Performance by terrain complexity

Figure 5 presents a Taylor diagram (Taylor, 2001) summarising the normalised standard deviation and correlation coefficient for three model groups, namely WRF-d01-2way, WRF-d02-2way, and MPAS, stratified by terrain complexity class. The one-way configurations are excluded from this figure as they are examined separately in the nesting comparison of Sect. 3.5. The two-way configurations represent the best-performing setup within each resolution level and thus provide a cleaner basis for the terrain-stratified comparison. The diagram reveals a systematic degradation of correlation with increasing terrain complexity across all model groups. The progressive reduction in MPAS bias relative to WRF configurations as terrain complexity increases is further documented in Table 5.

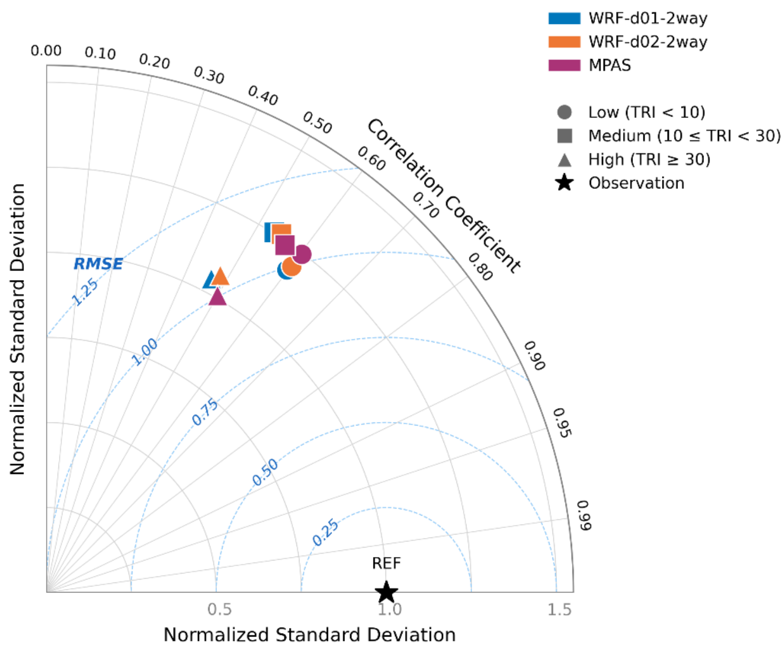


**Figure 5: Taylor diagram of normalised model performance stratified by terrain complexity. Colours indicate model groups; marker shapes indicate terrain complexity class (circle: low TRI; square: medium TRI; triangle: high TRI).**

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


In the low-complexity class, all three model groups perform comparably, with correlation coefficients ranging from 0.598 (WRF-d01-2way) to 0.603 (WRF-d02-2way and MPAS) and RMSE values between 2.11 and 2.16 m s−1 (Table 5). The points cluster tightly in the Taylor diagram, indicating that discretisation strategy has limited influence on forecast skill where topographic forcing is weak. Normalised standard deviations across all models slightly exceed unity (1.18–1.25), reflecting a modest overestimation of wind speed variability consistent with the positive bias noted in Sect. 3.1.

Model performances diverge more clearly in the medium and high complexity classes. For medium-complexity stations, MPAS achieves a CC of 0.567 and RMSE of 2.05 m s−1, compared to 0.535 and 2.18 m s−1 for WRF-d01-2way, and 0.549 and 2.17 m s−1 for WRF-d02-2way (Table 5). The difference is most pronounced at high-complexity sites, where MPAS records a CC of 0.500 and RMSE of 1.99 m s−1 against 0.466 and 2.08 m s−1 for WRF-d01-2way and 0.481 and 2.12 m s−1 for WRF-d02-2way (Table 5). In the Taylor diagram, the high-complexity points (triangles) shift noticeably further from the reference across all models, aligned with the expected degradation of forecast skill with increasing terrain ruggedness (Jiménez and Dudhia, 2012; Gómez-Navarro et al., 2015). The angular separation between MPAS and the WRF configurations is largest at high-complexity sites, which may reflect the absence of a discrete lateral boundary in MPAS's variable-resolution mesh and the associated reduction in interpolation errors at resolution transitions (Park et al., 2014; Harris and Durran, 2010). The normalised standard deviation also exhibits a terrain-dependent pattern: at high-complexity sites, MPAS (1.008) approaches the observed variability more closely than WRF-d01-2way (1.042) and WRF-d02-2way (1.064), indicating that MPAS not only reduces systematic bias but also better reproduces the amplitude of wind speed fluctuations in complex terrain.

**Table 5: Verification statistics stratified by terrain complexity class for the two-way WRF configurations and MPAS. ME: mean error; RMSE: root mean square error; CC: Pearson correlation coefficient. Statistics are aggregated over lead times of 6–47 h for January and June 2023 combined.**

| TRI Class | Model | ME (m s$^{-1}$) | RMSE (m s$^{-1}$) | CC |
|---|---|---|---|---|
| Low (n=159) | WRF-d01-2way | 1.02 | 2.11 | 0.598 |
| | WRF-d02-2way | 1.05 | 2.14 | 0.601 |
| | MPAS | 1.00 | 2.16 | 0.603 |
| Medium (n=196) | WRF-d01-2way | 1.07 | 2.18 | 0.535 |
| | WRF-d02-2way | 1.12 | 2.17 | 0.549 |
| | MPAS | 0.88 | 2.05 | 0.567 |
| High (n=154) | WRF-d01-2way | 0.65 | 2.08 | 0.466 |
| | WRF-d02-2way | 0.80 | 2.12 | 0.481 |
| | MPAS | 0.47 | 1.99 | 0.500 |

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


### 3.3 Spatial patterns of model performance

Figure 6 shows the spatial distribution of the better-performing model in terms of RMSE at each of the 509 evaluation stations, overlaid on the Terrain Ruggedness Index (TRI) background. MPAS achieves lower RMSE at 330 stations (64.8%) and WRF-d02-2way at 179 stations (35.2 %) of the 509 in total. While MPAS dominates across most of the domain, the spatial distribution of WRF 4 km's advantage is geographically structured and closely associated with coastal proximity rather than terrain complexity alone.

In the interior of the domain, MPAS outperforms WRF-d02-2way at 230 of 314 inland stations (73.2 %). This pattern is broadly consistent with the TRI-stratified results of Sect. 3.2, where MPAS's performance advantage over WRF configurations grows with increasing terrain complexity.

The coastal zones present a markedly different picture. Along the Aegean coast, WRF-d02-2way achieves lower RMSE at 44 of 63 stations (70 %), making it the dominant model in this sub-region. A similar, though less pronounced, reversal is observed along the Marmara coast and sections of the Black Sea coast. Along the Mediterranean coast the pattern is more mixed, with the WRF-d02-2way advantage concentrated in the southern coastal lowlands.

The physical mechanism underlying the coastal reversal is addressed in Sect. 3.6. As shown there, MPAS's aggregate ME advantage partly arises from a compensating bias structure in which it performs better at low-to-moderate wind speeds but exhibits stronger underprediction at high wind speeds. The monthly mean wind speed distributions shown in Fig. 7 confirm that coastal regions, particularly along the Aegean, are characterised by relatively higher mean wind speeds compared to inland areas in both January and June. This means the wind speed regime along the coast falls more frequently into the range where MPAS's underprediction tendency is most pronounced, which explains why WRF-d02-2way performs comparably or better in these areas without requiring an appeal to differences in grid geometry or coastline representation.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


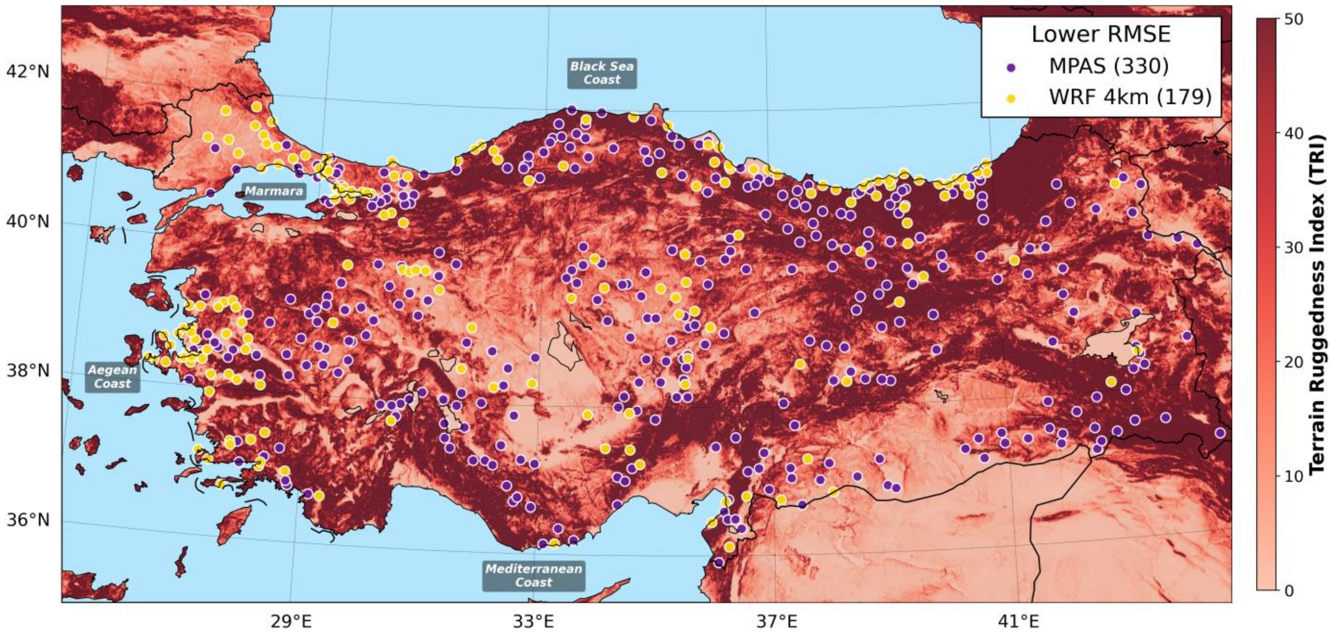


**Figure 6: Spatial distribution of the better-performing model in terms of RMSE at each of the 509 evaluation stations, overlaid on the Terrain Ruggedness Index (TRI). Purple and yellow markers indicate stations where MPAS and WRF 4 km (2-way) achieve lower RMSE, respectively.**

### 3.4 Temporal performance

Figure 7 presents the monthly mean spatial distribution of 10 m wind speed forecasts from MPAS and WRF-d02-2way, derived from lead times of 6–29 h to ensure that an equal number of simulation days contributes to each forecast hour. Both models exhibit regionally consistent seasonal patterns in mean wind speed between January and June. This seasonality is also evident in station observations, with the summer–winter asymmetry over inland regions attributed to thermally driven instability associated with rapid surface heating during summer months and strengthening pressure gradients. This effect is particularly pronounced over open and flat plateaus where surface friction is low (Şahin and Türkeş, 2013).

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


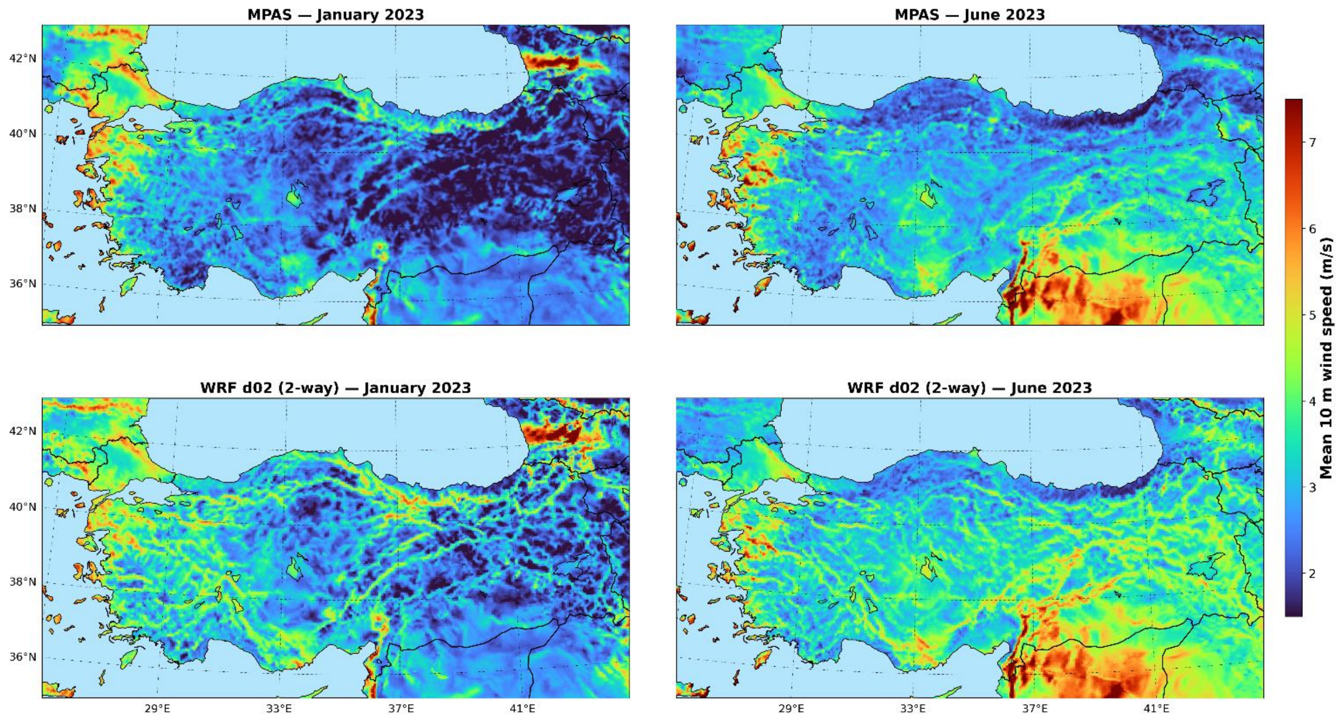


**Figure 7: Mean 10 m wind speed (m s−1) for January 2023 (left) and June 2023 (right) from MPAS (top) and WRF-d02-2way (bottom), averaged over the lead times (6–29 h) and all simulation days within each month.**

Figure 8 shows the corresponding Taylor diagrams for the two evaluation months, while Fig. 9 presents the diurnal cycle of ME and MAE averaged across all 509 stations for January and June 2023. The Taylor diagram (Fig. 8) demonstrates a seasonal contrast in model skill. All five configurations show higher correlation coefficients and lower normalised RMSE in June (CC: 0.46–0.56) compared to January (CC: 0.42–0.46), with the June points clustering closer to the reference in the diagram. Although the analysis is limited to two months of a single year, the higher model skill observed in June relative to January is consistent with the more stationary atmospheric regime that characterises summer over Türkiye. As synoptic forcing weakens during summer months, near-surface winds are increasingly governed by persistent thermal circulations associated with land surface heating (Şahin and Türkeş, 2013), providing a more stable and consistent signal for numerical models. In contrast, the winter regime over the Eastern Mediterranean is characterised by the interplay of multiple large-scale systems including the Siberian anticyclone, Mediterranean cyclones, and teleconnection patterns such as the NAO, ENSO, and EA/WR (Bağcacı et al., 2024), which introduce greater atmospheric variability and thus reduce model skill. Notably, MPAS achieves the highest CC in both months (January: 0.459; June: 0.557) and the lowest ME (January: 0.72 m s−1; June: 0.79 m s−1), maintaining its advantage over WRF configurations across both seasons.

The diurnal ME cycle (Fig. 9) reveals a pronounced hour-dependent bias structure that differs markedly between the two months. In January, WRF-d01-1way exhibits a strong nocturnal peak (ME ~1.75 m s−1 around 00–06 UTC) followed by a sharp reduction toward midday (~0.65 m s−1 around 13–14 UTC), before increasing to high values in the evening. The

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


particularly pronounced nocturnal bias in WRF-d01-1way, relative to the other configurations, is consistent with the compounding effect of two factors. First, the absence of two-way feedback prevents fine-scale boundary layer corrections resolved in the inner domain from propagating to the coarse parent (Harris and Durran, 2010). Second, at 12 km resolution, the heterogeneous surface conditions that govern stable nocturnal boundary layer structure are insufficiently resolved, a known limitation of kilometre-scale NWP under stable stratification (Jiménez and Dudhia, 2012). This diurnal pattern is broadly consistent with known difficulties of kilometre-scale NWP in representing stable nocturnal boundary layer conditions during winter, where surface-based inversions tend to suppress near-surface winds and models have been shown to overestimate turbulent mixing (Jiménez and Dudhia, 2012). The remaining configurations show a weaker but qualitatively similar pattern, with MPAS consistently maintaining the lowest ME throughout the day.

The MAE diurnal cycle (Fig. 9) follows a broadly similar pattern to that of ME but with reduced amplitude, as MAE measures the absolute magnitude of errors regardless of sign. In January, all models exhibit relatively flat MAE profiles (~1.6–1.8 m s−1) with the exception of WRF-d01-1way, whose MAE exceeds 2.2 m s−1 throughout the day. In June, a clear midday maximum is visible across all models, peaking around 13–15 UTC and coinciding with the period of strongest convective boundary layer development. The convergence of the four WRF configurations toward similar MAE values in June contrasts with their larger spread in January, suggesting that winter synoptic variability differentiates model performance more than summer thermally driven conditions.

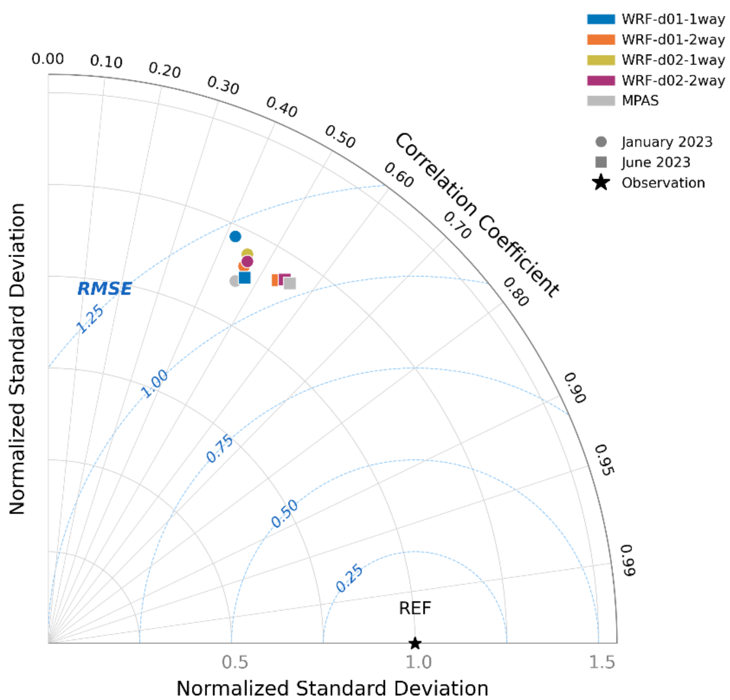


**Figure 8: Normalised Taylor diagram comparing model skill for January 2023 (circles) and June 2023 (squares). Colours denote model configurations. Each point represents the normalised standard deviation and correlation coefficient of hourly 10 m wind speed forecasts aggregated over 509 stations and all valid lead times (6–47 h).**

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


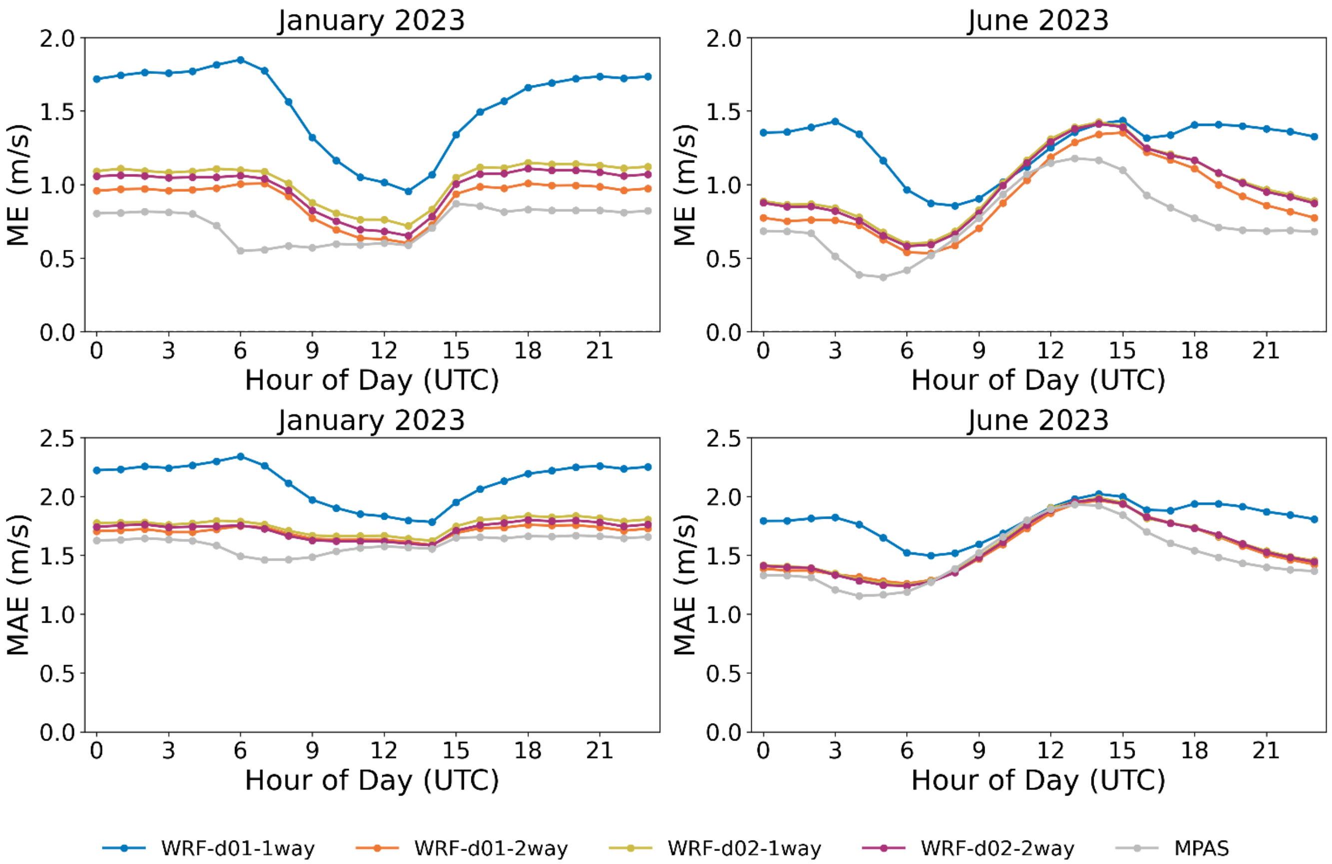


**Figure 9: Diurnal cycle of mean error (ME, top panels) and mean absolute error (MAE, bottom panels) averaged across 509 TSMS stations for January 2023 (left) and June 2023 (right). Values are computed from hourly forecasts aggregated over all lead times (6–47 h) and all stations.**

### 3.5 WRF nesting comparison

Table 6 presents ME, RMSE, and CC for the four WRF configurations across January and June 2023 and the three terrain complexity classes, enabling a systematic assessment of two distinct effects: the benefit of two-way feedback over one-way nesting, and the added value of increasing horizontal resolution from 12 km to 4 km.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


**Table 6: Performance metrics (ME, RMSE, CC) for WRF nesting configurations stratified by season and terrain complexity class. Values are aggregated over 509 TSMS stations and lead times of 6–47 h. ME and RMSE are given in m s−1.**

| **Configuration** | **Time Period** | | | | | | **TRI Class** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | January-2023 | | | June-2023 | | | Low TRI (n=159) | | | Medium TRI (n=196) | | | High TRI (n=154) | | |
| | ME | RMSE | CC | ME | RMSE | CC | ME | RMSE | CC | ME | RMSE | CC | ME | RMSE | CC |
| WRF-d01-1way | 1.51 | 2.80 | 0.418 | 1.24 | 2.35 | 0.473 | 1.30 | 2.36 | 0.536 | 1.36 | 2.35 | 0.476 | 1.04 | 2.31 | 0.403 |
| WRF-d01-2way | 0.89 | 2.34 | 0.461 | 0.93 | 2.13 | 0.534 | 1.02 | 2.11 | 0.598 | 1.07 | 2.18 | 0.535 | 0.65 | 2.08 | 0.466 |
| WRF-d02-1way | 1.01 | 2.43 | 0.455 | 1.02 | 2.16 | 0.546 | 1.07 | 2.15 | 0.603 | 1.13 | 2.18 | 0.550 | 0.82 | 2.13 | 0.481 |
| WRF-d02-2way | 0.97 | 2.37 | 0.462 | 1.01 | 2.15 | 0.545 | 1.05 | 2.14 | 0.601 | 1.12 | 2.17 | 0.549 | 0.80 | 2.12 | 0.481 |

Except for a marginal CC difference for the 4 km configurations, two-way nesting outperforms the one-way nesting strategy across metrics and stratifications within the same domain. At 12 km, WRF-d01-2way achieves lower ME than WRF-d01-1way, recording 0.89 m s−1 versus 1.51 m s−1 in January and 0.93 m s−1 versus 1.24 m s−1 in June, corresponding to reductions of 41 % and 25 % respectively. RMSE improvements are similarly pronounced, decreasing from 2.80 to 2.34 m s−1 in January and from 2.35 to 2.13 m s−1 in June (Table 6). The effect is smaller but consistent at 4 km, where two-way feedback reduces ME from 1.01 to 0.97 m s−1 in January and from 1.02 to 1.01 m s−1 in June, suggesting diminishing returns as the resolution of the inner domain increases. The asymmetry between the 12 km and 4 km two-way improvements suggests that the coarse parent domain is the primary beneficiary of the feedback mechanism. In two-way nesting, fine-scale flow patterns resolved by the inner domain are fed back to the parent, correcting its solution in ways that one-way nesting cannot achieve. Although these corrections are subsequently returned to the inner domain at the following time step, the incremental adjustments at 4 km resolution are small relative to the dominant synoptic-scale forcing, resulting in negligible differences between WRF-d02-1way and WRF-d02-2way performance metrics. When stratified by terrain complexity, the two-way improvement at 12 km is most pronounced at high-complexity sites, where ME drops from 1.04 to 0.65 m s−1 (Table 6), consistent with the expectation that inner-domain corrections are largest where topographic forcing on near-surface flow is strongest.

The comparison between WRF-d01-2way and WRF-d02-2way reveals a counterintuitive result: increasing horizontal resolution does not lead to consistent improvement. In January, WRF-d01-2way records lower ME (0.89 m s−1) and RMSE (2.34 m s−1) than WRF-d02-2way (ME: 0.97 m s−1; RMSE: 2.37 m s−1). In June the difference narrows, with ME values of 0.93 and 1.01 m s−1 respectively, but WRF-d01-2way still shows lower bias (Table 6). This pattern holds across all TRI classes. At high-complexity sites, WRF-d01-2way records ME of 0.65 m s−1 compared to 0.80 m s−1 for WRF-d02-2way, while RMSE values are similar (2.08 vs. 2.12 m s−1). The only exception is the correlation coefficient, where WRF-d02-2way performs slightly better in June (CC: 0.545 vs. 0.534), indicating that the finer domain captures temporal variability somewhat more accurately despite its higher mean bias.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


This result suggests that resolution increase alone is not sufficient to improve near-surface wind bias, and that the interaction between finer terrain representation and boundary layer parameterisation at kilometre scales may introduce compensating errors. At 4 km, WRF resolves topographic gradients more explicitly, but this may also amplify terrain-induced errors in the boundary layer scheme if sub-grid drag effects are not accounted for. In addition, the fine domain solution is constrained by lateral boundary conditions from the parent domain, which limits the practical benefit of the higher resolution (Warner et al., 1997; Harris and Durran, 2010). The physical mechanisms behind this behaviour are discussed further in Sect. 4.

### 3.6 Wind speed distribution analysis

Figure 10 presents the mean error as a function of observed wind speed bin for the three model groups across low-, medium-, and high-TRI classes. The results reveal a consistent and physically interpretable bias structure that is common to all models but differs in magnitude and wind speed threshold across configurations and terrain classes.

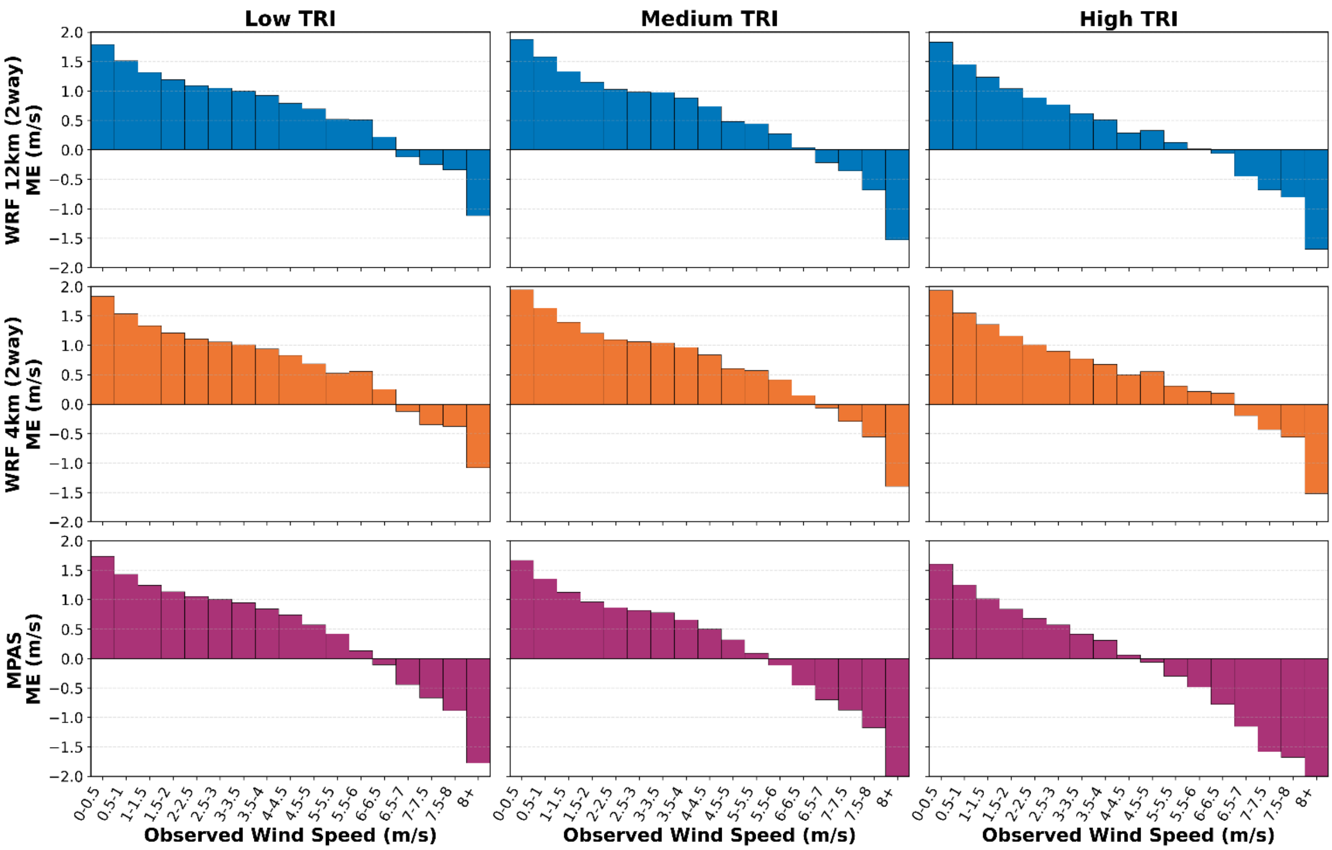


**Figure 10: Mean error (ME) as a function of observed wind speed bin for WRF 12 km (2-way), WRF 4 km (2-way), and MPAS across low (left), medium (centre), and high (right) terrain complexity classes. Each bar represents the mean ME of all station–hour pairs within the corresponding observed wind speed bin. Statistics are aggregated over 509 TSMS stations and lead times of 6–47 h for January and June 2023 combined.**

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


All three model groups exhibit a monotonically decreasing ME with increasing observed wind speed, starting from large positive bias at calm conditions (0–1 m s−1), reducing progressively through moderate speeds, and eventually shifting to negative bias at high wind speeds. This pattern is a well-documented characteristic of kilometre-scale NWP over complex terrain (Jiménez and Dudhia, 2012), reflecting the difficulty of capturing the full range of near-surface wind variability across different stability regimes.

The observed wind speed at which bias changes sign from positive to negative, referred to here as the zero-crossing threshold, provides a compact summary of each model's bias characteristics. For WRF-d01-2way and WRF-d02-2way, this threshold occurs consistently around 6.75 m s−1 across all three TRI classes. MPAS crosses zero earlier, at approximately 6.25 m s−1 in low TRI, 5.75 m s−1 in medium TRI, and 4.75 m s−1 in high TRI. This earlier transition reflects MPAS's lower overall positive bias documented in Sect. 3.1, but also indicates that MPAS begins to underpredict at comparatively moderate wind speeds in complex terrain.

The high-wind underprediction is most pronounced for MPAS at high-complexity sites, where ME reaches −3.28 m s−1 in the highest wind speed bin, compared to −1.69 m s−1 for WRF-d01-2way and −1.52 m s−1 for WRF-d02-2way. This indicates that while MPAS reduces the low-to-moderate wind overprediction that dominates the overall ME, it does so partly at the cost of underestimating strong wind events in rugged terrain. Whether this reflects a limitation of MPAS's surface layer scheme or a consequence of its mesh geometry near steep orography remains an open question.

Taken together, these results suggest that the aggregate ME advantage of MPAS over WRF configurations (Sect. 3.1) partially arises from a compensating effect: lower overprediction at calm conditions offsets stronger underprediction at high wind speeds. This bias decomposition also helps explain the spatial performance patterns observed in Fig. 6. In inland regions where mean wind speeds are generally lower, MPAS benefits from its reduced calm-condition bias, whereas along coastal areas where higher wind speeds are more frequent, the stronger high-wind underprediction of MPAS works against it, allowing WRF-d02-2way to perform comparably or better. These findings indicate that a blanket assertion of MPAS superiority over WRF is not warranted; the relative performance of the two modelling approaches depends on the local wind climate and the application in question.

## 4 Discussion

The results of this study show that MPAS generally outperforms WRF configurations for near-surface wind speed forecasting across Türkiye, with the performance gap widening with terrain complexity. This section examines the physical and numerical mechanisms that may underlie these findings, addresses the unexpected behaviour of WRF's nested configurations, and notes the limitations that constrain the interpretation of the results.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


### 4.1 MPAS advantage and its terrain dependence

MPAS achieves lower ME and RMSE than all WRF configurations across most stations (64.8 %) and all TRI classes, but the magnitude of the advantage depends strongly on terrain complexity. In low-complexity terrain, the three model groups perform comparably, with RMSE and ME differences within 0.05 m s−1, indicating that discretisation strategy has limited influence where topographic forcing is weak and the boundary layer is predominantly governed by large-scale synoptic flow. As terrain complexity increases, the advantage of MPAS grows: at high-complexity sites MPAS achieves ME of 0.47 m s−1 against 0.80 m s−1 for WRF-d02-2way and 0.65 m s−1 for WRF-d01-2way, corresponding to reductions of 41 % and 28 % respectively. One explanation for this terrain-dependent advantage lies in the absence of lateral nesting boundaries in MPAS. In WRF's nested configurations, the inner domain receives boundary conditions from the parent domain at fixed lateral interfaces, which can introduce spurious reflections and interpolation errors that propagate into the fine domain interior (Warner et al., 1997; Harris and Durran, 2010). These effects are most consequential in regions of complex terrain, where terrain-forced circulations interact with the nest boundary in ways that are difficult to represent accurately. MPAS's unstructured variable-resolution mesh achieves the resolution transition smoothly and continuously, eliminating the discrete boundary interface and its associated errors. This structural advantage appears most clearly at high-TRI stations where boundary layer and orography interactions are strongest.

### 4.2 Resolution increase and nesting feedback

A counterintuitive finding of this study is that increasing horizontal resolution from 12 km to 4 km within the WRF framework does not consistently improve near-surface wind forecasts. WRF-d01-2way outperforms WRF-d02-2way in terms of ME across both months and all TRI classes, despite the latter resolving topographic gradients at three times the spatial detail. Interpreting this result requires some care, however, since the two configurations differ not only in resolution but also in the role of the feedback mechanism. In two-way nesting, the inner domain corrections propagate back to the parent and improve its solution, whereas the inner domain itself benefits only indirectly at the following time step. The relatively weak improvement from one-way to two-way nesting at 4 km, documented in Sect. 3.5, suggests that the inner domain solution is not substantially altered by the feedback and that its performance is therefore more strongly constrained by the parent domain boundary conditions than by its own resolution.

It is worth noting that WRF-d02-2way marginally outperforms WRF-d01-2way in terms of correlation coefficient in June (CC: 0.545 vs. 0.534), suggesting that the finer domain captures some temporal variability more accurately even when its mean bias is higher. The resolution increase therefore does not uniformly degrade performance but fails to deliver consistent improvement across all metrics. This result echoes findings from several kilometre-scale downscaling studies over complex terrain (Jiménez and Dudhia, 2012; Gómez-Navarro et al., 2015) and points to a broader limitation: the added value of resolution is conditional on the ability of the physical parameterisations to exploit that resolution. At 4 km, WRF's boundary layer scheme operates near the boundary between the fully parameterised and partially resolved turbulence regimes, a transition zone in which

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


conventional PBL schemes were not originally designed to operate and where their assumptions may no longer hold (Wyngaard, 2004; Shin and Hong, 2015). The resulting grey-zone behaviour may introduce additional errors that offset the benefits of finer terrain representation. Furthermore, the lateral boundary conditions imposed on the 4 km domain by the 12 km parent constrain the mesoscale solution and may prevent the inner domain from fully developing the independent circulation features that higher resolution would otherwise produce (Warner et al., 1997; Harris and Durran, 2010).

### 4.3 Bias structure across the wind speed spectrum

The wind speed bin analysis (Sect. 3.6) reveals that the overall ME advantage of MPAS partly arises from a compensating bias structure: MPAS reduces overprediction at low-to-moderate wind speeds (below ~5 m s−1 in high-complexity terrain) but exhibits stronger underprediction at high wind speeds, reaching ME of −3.28 m s−1 in the highest observed wind speed bin at high-TRI stations. This asymmetry suggests that MPAS's surface layer and boundary layer schemes may suppress near-surface wind speeds more aggressively under stable conditions while underestimating the intensity of dynamically forced high-wind events. The practical implication is that MPAS's lower aggregate ME does not translate uniformly across the wind speed spectrum, and users working with high-complexity sites should account for this high-speed underprediction in applications sensitive to strong wind events. The positive bias at low-to-moderate wind speeds, common to both models, is consistent with findings from Shin et al. (2025), who reported similar overprediction tendencies in regional MPAS and WRF over complex terrain in the Korean Peninsula.

### 4.4 Limitations

Finally, both models employ nearly identical physics parameterisations, which means that the performance differences reported here reflect primarily the effects of discretisation strategy and domain configuration rather than physical process representation. Nevertheless, other numerical differences between the two models may also contribute to the observed performance gaps. MPAS uses a height-based vertical coordinate with smoothed terrain-following surfaces (Klemp, 2011), whereas WRF employs a hybrid sigma-pressure coordinate; this difference in vertical coordinate formulation can influence the representation of small-scale atmospheric motion over complex terrain (Shin et al., 2025). Furthermore, MPAS solves the governing equations using a finite-volume framework on an unstructured C-grid with third- and fourth-order horizontal flux operators, while WRF employs a finite-difference approach with fifth-order horizontal flux operators (Skamarock et al., 2012; Shin et al., 2025). However, Shin et al. (2025) found that sensitivity experiments targeting these numerical differences did not produce substantial changes in near-surface wind speed, suggesting that their contribution to surface-level performance gaps may be limited relative to mesh geometry and domain configuration effects. Disentangling these contributions would require dedicated numerical sensitivity experiments beyond the scope of this study.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


## 5 Conclusions

This study evaluated short-term near-surface wind speed predictions from regional MPAS and nested WRF configurations against observations from 509 TSMS stations across Türkiye during January and June 2023. Five model configurations were compared, namely WRF-d01-1way, WRF-d01-2way, WRF-d02-1way, WRF-d02-2way, and MPAS with a variable-resolution mesh centred on Türkiye, using ME, MAE, RMSE, and CC stratified by terrain complexity via the Terrain Ruggedness Index. The following principal conclusions are drawn.

MPAS achieves the lowest error across all aggregate metrics, recording ME of 0.76 m s−1, RMSE of 2.16 m s−1, and CC of 0.541, compared to the best WRF configuration, WRF-d02-2way, with ME of 0.99 m s−1, RMSE of 2.27 m s−1, and CC of 0.536. All models exhibit systematic positive bias, consistent with the overprediction of near-surface wind speeds reported in kilometre-scale NWP studies over complex terrain.

The performance gap between MPAS and WRF is strongly terrain dependent. In low-complexity terrain, all configurations perform comparably. At high-complexity sites, MPAS reduces ME by 41 % relative to WRF-d02-2way and by 28 % relative to WRF-d01-2way, pointing to the structural advantage of MPAS's smooth mesh transition over WRF's discrete nesting boundaries.

Increasing WRF resolution from 12 km to 4 km does not consistently improve near-surface wind bias. WRF-d01-2way outperforms WRF-d02-2way in ME across both months and all TRI classes, suggesting that grey-zone boundary layer limitations and lateral boundary constraints offset the benefits of finer terrain representation. WRF-d02-2way does, however, marginally outperform WRF-d01-2way in correlation coefficient in June (CC: 0.545 vs. 0.534), indicating that the finer domain captures some temporal variability more accurately even when its mean bias is higher.

Two-way nesting provides meaningful improvement at 12 km, reducing ME by 41 % in January relative to WRF-d01-1way, but offers diminishing returns at 4 km. The coarse parent domain is the primary beneficiary of the feedback mechanism, while incremental corrections at 4 km resolution remain small relative to the dominant synoptic-scale forcing.

MPAS's aggregate ME advantage partly reflects a compensating bias structure: reduced overprediction at calm and moderate wind speeds is offset by stronger underprediction at high wind speeds, reaching ME of −3.28 m s−1 at the highest wind speed bin in high-complexity terrain. This wind speed dependence also explains the spatial reversal observed along the Aegean coast, where higher wind speeds are more frequent and WRF-d02-2way performs comparably or better than MPAS at 70 % of stations.

Taken together, these results show that MPAS's variable-resolution unstructured discretisation is a viable alternative to WRF's nested domain approach for regional near-surface wind forecasting over heterogeneous terrain. However, the relative performance of the two models depends on the local wind climate, and a blanket assertion of MPAS superiority is not warranted. Resolution increase alone is insufficient to improve wind forecasts without concurrent advances in boundary layer parameterisation for the grey zone.

## Code and data availability

Zenodo (Yalçın et al., 2026; https://doi.org/10.5281/zenodo.20606508). The package includes the WRF namelists, the MPAS namelist and stream files, the JIGSAW mesh-generation scripts, the TRI-processing code, the station-selection workflow, the model-to-station extraction scripts, the Python verification pipeline, and the configuration files required to reproduce the simulations and evaluation reported here. The archived code and data are distributed under the MIT License. The GFS forcing data are publicly available from the NCAR Research Data Archive (National Centers for Environmental Prediction et al., 2015). The TSMS station observations are subject to access restrictions by the Turkish State Meteorological Service; researchers may submit a data request through the TSMS data portal at https://mevbis.mgm.gov.tr/. The derived station-level verification tables and the scripts used to reproduce the reported figures and statistics are archived in the Zenodo repository (Yalcin et al., 2026).

## Author contributions

RDY: conceptualization, methodology, software, data curation, formal analysis, visualization, and original draft preparation. MTY: supervision, review, and editing. IY: supervision, review, and editing. OLS: review and editing. SO: review and editing. AUGS: review and editing. All authors have read and agreed to the submitted version of the manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Acknowledgements

The authors thank the Turkish State Meteorological Service (TSMS) for providing the station observation data used in this study. The numerical calculations reported in this paper were fully performed at TÜBİTAK ULAKBİM, High Performance and Grid Computing Center (TRUBA resources).

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


**References**

Bağçacı, S. Ç., Yücel, I., Yılmaz, M. T., Şen, O. L., and Ludwig, P.: High-resolution climate simulations over the Eastern Mediterranean–Black Sea region using the pseudo-global warming method with a CMIP6 ensemble, J. Geophys. Res. Atmos., 129, e2023JD040145, https://doi.org/10.1029/2023JD040145, 2024.

Barber, S., Nordborg, H., and Sütterlin, B.: Machine learning applications for improving wind power prediction in hilly terrain, Wind Energ. Sci., 7, 2231–2254, https://doi.org/10.5194/wes-7-2231-2022, 2022.

Chen, L., Li, G., Zhang, F., and Wang, C.: Simulation uncertainty of near-surface wind caused by boundary layer parameterization over the complex terrain, Front. Energy Res., 8, 554544, https://doi.org/10.3389/fenrg.2020.554544, 2020.

Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., and Feitosa, E.: A review on the young history of the wind power short-term prediction, Renew. Sustain. Energy Rev., 12, 1725–1744, https://doi.org/10.1016/j.rser.2007.01.015, 2008.

Çetin, I. I., Yücel, I., Yılmaz, M. T., and Önol, B.: Historical variability of Coupled Model Intercomparison Project Version 6 (CMIP6)-driven surface winds and global reanalysis data for the Eastern Mediterranean, Theor. Appl. Climatol., 155, 4101–4121, https://doi.org/10.1007/s00704-024-04869-y, 2024.

Davies, H. C.: A lateral boundary formulation for multi-level prediction models, Q. J. Roy. Meteor. Soc., 102, 405–418, https://doi.org/10.1002/qj.49710243210, 1976.

Du, S. and Stechmann, S. N.: A universal predictor-corrector approach for minimizing artifacts due to mesh refinement, J. Adv. Model. Earth Sy., 15, e2023MS003688, https://doi.org/10.1029/2023MS003688, 2023.

Engwirda, D.: JIGSAW-GEO (1.0): locally orthogonal staggered unstructured grid generation for general circulation modelling on the sphere, Geosci. Model Dev., 10, 2117–2140, https://doi.org/10.5194/gmd-10-2117-2017, 2017.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Rev. Geophys., 45, RG2004, https://doi.org/10.1029/2005RG000183, 2007.

Foley, A. M., Leahy, P. G., Marvuglia, A., and McKeogh, E. J.: Current methods and advances in forecasting of wind power generation, Renew. Energy, 37, 1–8, https://doi.org/10.1016/j.renene.2011.05.033, 2012.

Gómez-Navarro, J. J., Raible, C. C., and Dierer, S.: Sensitivity of the WRF model to PBL parametrisations and nesting techniques: evaluation of wind storms over complex terrain, Geosci. Model Dev., 8, 3349–3363, https://doi.org/10.5194/gmd-8-3349-2015, 2015.

Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, Atmos. Chem. Phys., 14, 5233–5250, https://doi.org/10.5194/acp-14-5233-2014, 2014.

Gürer, K., Zhao, Z., Cai, C., and Avise, J. C.: The evaluation of global and regional applications of the Model for Prediction Across Scales-Atmosphere (MPAS) against the Weather Research and Forecasting (WRF) model over California for a winter (2013 DISCOVER-AQ) and summer (2016 CABOTS) episode, Atmosphere, 15, 1248, https://doi.org/10.3390/atmos15101248, 2024.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


Hagos, S., Leung, R., Rauscher, S. A., and Ringler, T.: Error characteristics of two grid refinement approaches in aquaplanet simulations: MPAS-A and WRF, Mon. Weather Rev., 141, 3022–3036, https://doi.org/10.1175/MWR-D-12-00338.1, 2013.

Hanifi, S., Liu, X., Lin, Z., and Lotfian, S.: A critical review of wind power forecasting methods – past, present and future, Energies, 13, 3764, https://doi.org/10.3390/en13153764, 2020.

Harris, L. M. and Durran, D. R.: An idealized comparison of one-way and two-way grid nesting, Mon. Weather Rev., 138, 2174–2187, https://doi.org/10.1175/2010MWR3080.1, 2010.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models, J. Geophys. Res., 113, D13103, https://doi.org/10.1029/2008JD009944, 2008.

Jiménez, P. A. and Dudhia, J.: Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model, J. Appl. Meteorol. Clim., 51, 300–316, https://doi.org/10.1175/JAMC-D-11-084.1, 2012.

Jolliffe, I. T. and Stephenson, D. B. (Eds.): Forecast Verification: A Practitioner's Guide in Atmospheric Science, 2nd Edn., Wiley-Blackwell, Oxford, UK, https://doi.org/10.1002/9781119960003, 2012.

Klemp, J. B.: A terrain-following coordinate with smoothed coordinate surfaces, Mon. Weather Rev., 139, 2163–2169, https://doi.org/10.1175/MWR-D-10-05046.1, 2011.

Kramer, M., Heinzeller, D., Hartmann, H., van den Berg, W., and Steeneveld, G.-J.: Assessment of MPAS variable resolution simulations in the grey-zone of convection against WRF model results and observations, Clim. Dynam., 55, 253–276, https://doi.org/10.1007/s00382-018-4562-z, 2020.

Murphy, A. H. and Winkler, R. L.: A general framework for forecast verification, Mon. Weather Rev., 115, 1330–1338, https://doi.org/10.1175/1520-0493(1987)115<1330:AGFFFV>2.0.CO;2, 1987.

Nakanishi, M. and Niino, H.: Development of an improved turbulence closure model for the atmospheric boundary layer, J. Meteorol. Soc. Jpn., 87, 895–912, https://doi.org/10.2151/jmsj.87.895, 2009.

NASA JPL: NASA Shuttle Radar Topography Mission Global 3 arc second, NASA Land Processes Distributed Active Archive Center [data set], https://doi.org/10.5067/MEASURES/SRTM/SRTMGL3.003, 2013.

National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce: NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive, Research Data Archive, NSF NCAR [data set], https://doi.org/10.5065/D65D8PWK, 2015.

Park, S.-H., Klemp, J. B., and Skamarock, W. C.: A comparison of mesh refinement in the global MPAS-A and WRF models using an idealized normal-mode baroclinic wave simulation, Mon. Weather Rev., 142, 3614–3634, https://doi.org/10.1175/MWR-D-14-00004.1, 2014.

Riley, S. J., DeGloria, S. D., and Elliot, R.: A terrain ruggedness index that quantifies topographic heterogeneity, Intermountain J. Sci., 5, 23–27, 1999.

Shin, H. H. and Hong, S.-Y.: Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions, Mon. Weather Rev., 143, 250–271, https://doi.org/10.1175/MWR-D-14-00116.1, 2015.

https://doi.org/10.5194/egusphere-2026-3355
Preprint. Discussion started: 20 July 2026


Shin, U., Park, S.-H., Skamarock, W., Lee, G., Lim, K.-S., and Joh, M.: Comparison of numerical characteristics in the regional Model for Prediction Across Scales (MPAS) and Weather Research and Forecasting (WRF) models using low-level wind analysis during heavy snowfall episodes over complex terrain, Mon. Weather Rev., 153, 471–490, https://doi.org/10.1175/MWR-D-24-0053.1, 2025.

Skamarock, W. C., Klemp, J. B., Duda, M. G., Fowler, L. D., Park, S.-H., and Ringler, T. D.: A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tessellations and C-grid staggering, Mon. Weather Rev., 140, 3090–3105, https://doi.org/10.1175/MWR-D-11-00215.1, 2012.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A description of the Advanced Research WRF Model Version 4, NCAR Tech. Note NCAR/TN-556+STR, National Center for Atmospheric Research, Boulder, CO, USA, https://doi.org/10.5065/1DFH-6P97, 2019.

Şahin, S. and Türkeş, M.: Contemporary surface wind climatology of Turkey, Theor. Appl. Climatol., 113, 337–349, https://doi.org/10.1007/s00704-012-0789-5, 2013.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, https://doi.org/10.1029/2000JD900719, 2001.

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M. A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., and Cuenca, R. H.: Implementation and verification of the unified Noah land surface model in the WRF model, in: 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction, Seattle, WA, USA, Paper 11.2, American Meteorological Society, 2004.

Thompson, G. and Eidhammer, T.: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone, J. Atmos. Sci., 71, 3636–3658, https://doi.org/10.1175/JAS-D-13-0305.1, 2014.

Warner, T. T., Peterson, R. A., and Treadon, R. E.: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction, B. Am. Meteorol. Soc., 78, 2599–2617, https://doi.org/10.1175/1520-0477(1997)078<2599:ATOLBC>2.0.CO;2, 1997.

Whiteman, C. D.: Mountain Meteorology: Fundamentals and Applications, Oxford University Press, New York, USA, 355 pp., ISBN 9780195132717, 2000.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, 3rd Edn., Academic Press, Oxford, UK, https://doi.org/10.1016/C2017-0-03921-6, 2011.

Wyngaard, J. C.: Toward numerical modeling in the "terra incognita", J. Atmos. Sci., 61, 1816–1826, https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2, 2004.

Yalçın, R. D., Yilmaz, M. T., Yucel, I., Sen, O. L., Oruc, S., and Şenocak, A. U. G.: Code and data for: Near-surface wind prediction over terrain of varying complexity: a comparison of MPAS-Atmosphere and nested WRF, Zenodo [code and data set], https://doi.org/10.5281/zenodo.20606508, 2026.