A spatial evaluation of high-resolution wind fields from empirical and dynamical modeling in hilly and mountainous terrain

Christoph Schlager[1], Gottfried Kirchengast[1,2], Juergen Fuchsberger[1], Alexander Kann[3], and Heimo Truhetz[1]

[1]Wegener Center for Climate and Global Change (WEGC), University of Graz, Graz, Austria.
[2]Institute for Geophysics, Astrophysics, and Meteorology/Institute of Physics, University of Graz, Graz, Austria.
[3]Department of Forecasting Models, Central Institute for Meteorology and Geodynamics (ZAMG), Vienna, Austria.

Correspondence: Christoph Schlager (christoph.schlager@uni-graz.at)

Abstract. Empirical high-resolution surface wind fields, automatically generated by a weather diagnostic application, the WegenerNet Wind Product Generator (WPG), were intercompared with wind field analysis data from the Integrated Nowcasting through Comprehensive Analysis (INCA) system and with dynamical climate model wind field data from the non-hydrostatic climate model COSMO-CLM. The INCA analysis fields are available at a horizontal grid spacing of 1 km × 1 km, whereas the COSMO model fields are from simulations at a 3 km × 3 km grid. The WPG, developed by Schlager et al. (2017, 2018), generates diagnostic fields at a high resolution grid of 100 m × 100 m, using observations from two dense meteorological station networks: The WegenerNet Feldbach Region (FBR) and its alpine sister network, the WegenerNet Johnsbachtal (JBT). The high-density WegenerNet FBR is located in southeastern Styria, Austria, a region predominated by a hilly terrain and small differences in altitude. The network consists of more than 150 meteorological stations. The WegenerNet JBT contains eleven meteorological stations at elevations ranging from about 600 m to 2200 m in a mountainous region in northern Styria.

The wind fields of these different empirical/dynamical modeling approaches were intercompared for thermally induced and strong wind events, using hourly temporal resolutions as supplied by the WPG, with the focus on evaluating spatial differences and displacements between the different datasets. For this comparison, a novel neighborhood-based spatial wind verification methodology based on fractions skill socres (FSS) is used to estimate the modeling performances. All comparisons show an increasing FSS with increasing neighborhood size. In general, the spatial verification indicates a better statistical agreement for the hilly WegenerNet FBR than for the mountainous WegenerNet JBT. The results for the WegenerNet FBR show a better agreement between INCA and WegenerNet than between COSMO and WegenerNet wind fields, especially for large scales (neighborhoods). In particular, COSMO-CLM clearly underperforms in case of thermally induced wind events. For the JBT region, all spatial comparisons indicate little overlap at small neighborhood sizes and in general large biases of wind vectors occur between the dynamical (COSMO) and analysis (INCA) fields and the diagnostic (WegenerNet) reference dataset.

Furthermore, gridpoint-based error measures were calculated for the same evaluation cases. The statistical agreement, estimated for the vector-mean wind speed and wind directions show again a better agreement for the WegenerNet FBR than for the WegenerNet JBT region. In general, the difference between modeled and observed wind directions is smaller for strong wind speed events than for thermally induced ones. A combined examination of all spatial and gridpoint-based error measures shows that COSMO-CLM with its limited horizontal resolution of 3 km × 3 km and hence, a too smoothed orography, is not

able to represent small-scale wind patterns. The results for the JBT region indicate that the INCA analysis fields generally overestimate wind speeds in the summit regions. For strong wind speed events the wind speed in the valleys is underestimated by INCA, however. Regarding the WegenerNet diagnostic wind fields, the statistics show decent performance in the FBR and somewhat overestimated wind speeds for strong wind speed events in the Enns valley of the JBT region.

5   *Copyright statement.*

1   Introduction

Surface wind is often considered as one of the most difficult meteorological variables to model, particularly over areas of complex terrain like the Alps (Whiteman, 2000; Sfetsos, 2002; Abdel-Aal et al., 2009; Gómez-Navarro et al., 2015). Therefore a realistic wind field modeling with course-resolution models is not possible and moreover hampers a simple interpolation of
10   wind station data onto regular grids. Innovation in computer sciences, new methods in weather analysis or nowcasting models, advanced software architectures used in regional climate models (RCMs) and the growing power of computers meanwhile led to highly-resolved outputs from such models at the 1-km scale (Awan et al., 2011; Suklitsch et al., 2011; Prein et al., 2013b, 2015; Leutwyler et al., 2016; Kendon et al., 2017).

These models, however, contain various limitations and sources of uncertainties. In case of weather analysis fields, which
15   is a mixed empirical and dynamical modeling from data fusion, they result from too little meteorological station and remote sensing data, and in case of dynamical regional climate models (RCMs) they include deviations in the driving data set, physical and numerical approximations, as well as parameterizations of processes at the sub-grid scale (Gómez-Navarro et al., 2015).

To evaluate and improve these analysis and models, meteorological observations and especially gridded empirical datasets at high spatial and temporal resolutions are needed. The model outputs on their side generally represent the involved processes
20   as areal averages rather than on a point-scale (Osborn and Hulme, 1998; Prein et al., 2015). Therefore, gridded meteorological evaluation datasets, with each (aggregated) grid value being a best-estimate average of the grid cell observations, are the most appropriate evaluation datasets (Haylock et al., 2008; Haiden et al., 2011; Hiebl and Frei, 2016).

To investigate weather and climate on a local 1-km scale as well as evaluating RCMs, the Wegener Center at the University of Graz operates two high-resolution meteorological station networks (Fig. 1): the very high-density WegenerNet Feldbach
25   Region (FBR) in southeastern Styria, Austria (Fig. 1b) and the high-density WegenerNet Johnsbachtal (JBT) in northern Styria, Austria (Fig. 1c); details introduced in section 2 below.

In this study, we intercompare the empirical WegenerNet wind fields (Schlager et al., 2017, 2018) with empirical-dynamical wind field analysis data from the Integrated Nowcasting through Comprehensive Analyses (INCA) (Haiden et al., 2011) system and with dynamical climate model data from the Consortium for Small Scale Modeling (COSMO) Model in Climate
30   Mode (CLM) (Schättler et al., 2016). The intercomparisons aim at getting useful and robust information about performance limits for these empirical and dynamical modeling approaches for regions with very different topographic characteristics and

weather situations. Furthermore, we co-analyze the impact of different horizontal resolutions, which inevitably will always be a challenge for the wide diversity of data products typically available.

In order to provide more detailed information about the INCA modeling performance, we evaluate the INCA wind fields against the WPG wind fields from the WegenerNet FBR and the WegenerNet JBT for representative types of weather events.

5   To analyze the impact of the spatial resolution for the WegenerNet FBR, we use again the WPG wind fields as reference data for the evaluation of the COSMO-CLM wind fields. Due to missing overlapping COSMO-CLM and WegenerNet data from the same period in the WegenerNet JBT, the evaluation of the COSMO-CLM wind fields in this region is performed using INCA datasets.

Besides traditional gridpoint based verification methods, we use a novel wind verification methodology, recently developed
10 by Skok and Hladnik (2018). This neighborhood-based spatial verification method avoids the "double penalty" problem and can distinguish forecasts depending on the spatial displacement of wind patterns (Skok and Hladnik, 2018). So our primary motivation of this study is indeed to explore an provide improved insight, by careful intercomparisons the relative performance strength and weakness of empirical an dynamical wind field modeling at high-spatial resolution over complex terrain where actual wind station observations will generally be available at sparse station density.

15   The paper is structured as follows. Section 2 provides a description of the study areas and basic information about the model data. Section 3 presents defined evaluation cases and the methodology for the automatic selection of typical wind events followed by a description of the methods used to evaluate model results. In the following section 4 results are presented and discussed in detail. Finally, in Section 5 we summarize our results and draw our conclusions.

2   Study Areas and Model Data

20 ### 2.1   Study Areas

The first study area, the WegenerNet FBR (indicated by the lower white-filled rectangle in Fig. 1a, enlarged in b) lies in the Alpine foreland of southeastern Styria, Austria, centered near the town of Feldbach (46.93°N, 15.90°E). It covers a dense grid of 154 meteorological stations within an area of about 22 km × 16 km, in a hilly terrain, characterized by small differences in altitude (Kirchengast et al., 2014). The typical difference in altitude between the valleys and the crests is about 100 m and the
25 highest peak is the Gleichenberg Kogel, with an elevation of 598 m.

This region, with a more Alpine climate at the valley floors and more Mediterranean climate along hillsides is quite sensitive to climate change (Wakonigg, 1978; Kabas et al., 2011; Hohmann et al., 2018). Furthermore, it exhibits rich weather variability, especially through strong convective activity and severe weather in summer (Kirchengast et al., 2014; Kann et al., 2015a; O et al., 2017, 2018; Schroeer and Kirchengast, 2018). The wind fields in this study area are characterized by thermally induced
30 local flows and influenced by thermally-driven regional wind systems with weak wind speeds, caused by a dynamical process called Alpine pumping (Lugauer and Winkler, 2005). Results related to the WPG-diagnosed empirical wind fields in the WegenerNet FBR can be found in Schlager et al. (2017, 2018).

The second study area, the WegenerNet JBT (indicated by the upper white-filled rectangle in Fig. 1a, enlarged in c), is named after the Johnsbachtal river basin (location of village Johnsbach 47.54°N, 14.58°E) and situated in the eastern Alpine region, in the *Ennstaler Alps* and the *Gesäuse National Park*, in the northern Styria, Austria. The terrain of this mountainous region is characterized by large differences in elevation. The Hochtor, with an elevation of 2369 m, is the highest summit, and the

5 valleys are roughly at a height from 600 m to 800 m (Strasser et al., 2013; Schlager et al., 2018). This region spans an area of about 16 km ×17 km and comprises eleven irregularly distributed meteorological stations including two summit stations at altitudes of 2.191 m and 1.969 m (Schlager et al., 2018).

The climate is Alpine with mean annual temperatures of around 8 °C to 0 °C and an annual precipitation of about 1.500 mm to 1.800 mm from the valley to the summit regions (Wakonigg, 1978; Prettenthaler et al., 2010). Typical for this region are

10 thermally induced local flows and westerly-flow synoptic weather conditions. Details related to first studies and their results as well as to the cooperation and partnerships can be found in Strasser et al. (2013), and in most up-to-date form in Schlager et al. (2018). Recently, Schlager et al. (2018) computed and evaluated WPG-generated empirical wind fields in this region.

2.2 WegenerNet data

The data acquired from the two WegenerNet regions FBR and JBT are automatically quality controlled and processed by

15 the WegenerNet Processing System (WPS), consisting of four subsystems (Kirchengast et al., 2014): The Command Receive Archiving System transfers raw measurement data via wireless transmission to the WegenerNet database in Graz, the Quality Control System checks the data quality, the Data Product Generator (DPG) generates regular station time series and gridded fields of weather and climate products, and the Visualization and Information System offers the data to users via the Wegener-Net data portal (www.wegenernet.org).

20 Besides weather and climate time series, the DPG generates, based on a spatial interpolation of the station observations, gridded fields of the variables temperature, precipitation and relative humidity for the WegenerNet FBR. In case of temperature, lapse rates are estimated from the temperature observations at the numerous different station altitudes. These lapse rates are used for generating temperature fields over the hilly terrain. For temperature and relative humidity fields, inverse-distance weighted interpolation is used, and for precipitation fields, inverse-distance-squared weighted interpolation is used. The gridded

25 products of the WegenerNet FBR are available to users in near-real time with a latency of about 1-2 hours. Kirchengast et al. (2014) and Kabas (2012) provide detailed information about the subsystems of the WPG.

The DPG furthermore includes a newly developed wind field application, Wind Product Generator (WPG), which provides high-resolution wind fields for the WegenerNet FBR as well as for the WegenerNet JBT. The WPG uses the freely available empirical California Meteorological Model (CALMET) as core tool and generates, based on meteorological observations,

30 terrain elevations and information about land use, mean wind fields at 10 m and 50 m height levels with a spatial resolution of 100 m × 100 m and a temporal resolution of 30 minutes, again with a maximum latency of about 1-2 hours. For the WegenerNet FBR, the gridded wind fields are available starting in 2007 and for the WegenerNet JBT starting in 2012. A detailed description of the WPG application and the statistical results for the WegenerNet FBR can be found in Schlager et al. (2017). More information regarding statistical results related to the WegenerNet JBT as well as information regarding evaluation results

from five-year climate data of the WegenerNet JBT in comparison to nine-year climate data from the WegenerNet FBR can be found in Schlager et al. (2018).

2.3 INCA and COSMO-CLM data

The INCA system has been developed at the Central Institute for Meteorology and Geodynamics (ZAMG) in Vienna, Austria,
5 to provide realistic analyses and nowcasts of quantities of several meteorological variables for the highly mountainous and the overall complex terrain of Austria. In case of the variable wind, the system operationally generates spatially distributed analysis wind fields in 3D and for 10-m height above ground with a horizontal grid spacing of 1 km × 1 km and a temporal resolution of 1 hour.

The basic idea of the INCA wind module is to statistically correct a numerical weather prediction (NWP) model first guess
10 (i.e., in operational mode the latest available NWP model run) with latest observational data which are not entering the NWP data assimilation. Thus, the skill of the INCA analysis depends on the station density, their representativeness and the spatial distribution of station observations, as well as on the skill of the NWP model providing the first guess. The impact of the NWP model on the analysis skill is further discussed in Kann et al. (2015b).

In this study, NWP model outputs used as first guess are generated by the revised version of ARPEGE-ALADIN (ALARO)
15 (Wang et al., 2006) with a horizontal grid spacing of 4.8 km × 4.8 km. Therefore the output shows errors in regions with low station density (Prein, 2013; Haiden et al., 2011). The INCA wind fields have already been evaluated for a moderately hilly region in the north of Austria (47.78°N–49°N, 138°E–17°E), where the wind analyses shows significantly higher errors compared to the statistical results from other meteorological variables (Haiden et al., 2011).

Regarding the COSMO-CLM (Rockel et al., 2008), the climate version of the "Lokalmodell", we use available wind fields
20 generated with the model version 5.0. These wind fields were generated during the course of a previous study and are available for the period 2008 to 2010. The data have a comparatively coarse horizontal resolution of 3 km × 3 km on an hourly basis that is nevertheless the highest resolution proper available. This limited resolution leads to a smoothed orography, which may result in different wind patterns with errors in wind speed or direction. Furthermore, the winds may be displaced by an incorrect position of the topographic slopes (Skok and Hladnik, 2018; Prein et al., 2013b).

25 The numerical settings for the simulations are based on the *third order, two time-level, Runge–Kutta, split-explicit* scheme and the physical settings are based on shallow convection. In this model setup, deep convection is resolved explicitly, which means that the used simulations are referred to as convection permitting climate simulations (CPCSs) (Prein et al., 2013b). Furthermore, the simulations are driven by the integrated forecast system (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). These driving data are provided in a horizontal grid spacing of about 25 km at mid latitudes,
30 and are calculated for 91 vertical layers (Bechtold et al., 2008).

The coordinate system of the INCA and the COSMO datasets are transformed into WGS84 / UTM zone 33N coordinates. Furthermore, we resampled the wind fields from these two models onto the WegenerNet FBR and WegenerNet JBT grid, using a bilinear interpolation method. Based on extensive sensitivity tests regarding different interpolation methods, we concluded that the statistical results are not significantly dependent on the interpolation method.

3 Evaluation events and methods

3.1 Events for wind field evaluation

The WegenerNet, INCA and COSMO wind fields are intercompared for two representative types of wind events: thermally induced wind events and strong wind events. For this purpose, we defined eight evaluation cases, four for each of the two study areas (Table 1). For the cases shown in Table 1 we use the WegenerNet data as reference, expect for evaluating the COSMO-CLM wind fields for the WegenerNet JBT. The reason for this is, that the COSMO-CLM data used in this study are available from 01/2008-12/2009, but the WegenerNet JBT data are available only as of 01/2012, since this latter network was sufficiently completed for long term monitoring only since 2012 (Table 1, cases COSMOvsINCA_therm_JBT and COSMOvs-INCA_strong_JBT).

In both study areas, autochthonous weather conditions mainly lead to thermally induced wind systems, meaning that the wind fields are controlled by small-scale temperature and pressure gradients. These autochthounous days are characterized by small synoptic influences, cloudless or nearly cloudless skies, low relative humidity and increased radiation fluxes between the Earth surface and the atmosphere (Prettenthaler et al., 2010). Due to frequently occurring temperature inversions in relation to clear sky and high pressure weather conditions in winter, which often leads to a stable atmospheric stratification in the whole WegenerNet FBR and in the valley regions of the WegenerNet JBT, autochthonous days are only selected from spring, summer and fall (March to October).

The automatic selection of thermally induced and strong wind events is done based on thresholds, that we defined based on sound physical and careful sensitivity checks summarized in Table 2. For the estimation of autochthonous days, we compared the observed wind speed (v), relative humidity (rh), and net radiation (Q_n) from selected stations with the respective thresholds. A further criterion for the selection of such days is a high daily global radiation, which indicates fair weather conditions. For this purpose, we compared the daily mean modeled global radiation ($Q_{g,m}$) for clear sky conditions with the observed daily mean net radiation ($Q_{n,o}$) for the WegenerNet FBR and with the observed daily mean global radiation ($Q_{g,o}$) for the WegenerNet JBT at a defined station locations (Table 2, reference data). The reason for the comparison of $Q_{g,m}$ with $Q_{n,o}$ for the WegenerNet FBR is that this station network includes no global radiation sensors. Due to the almost linear relationship between the daytime $Q_{g,o}$ and $Q_{n,o}$ for clear sky conditions we find that the same selection method can robustly be applied to both study areas by defining different thresholds (Table 2, $Q_{g,m}$-$Q_{n(g),o}$).

The modeling of the global radiation is done based on ESRI's ArcGIS Area Solar Radiation Tool. This tool is designed for local landscape scales and derives the incoming solar radiation based on a digital elevation model. Small differences of daily mean values between $Q_{g,m}$ and $Q_{n(g),o}$ indicates fair weather conditions and high global radiations during the day. If all criteria are fulfilled for a given day, the data from that day are included into the thermally induced wind events dataset.

The strong wind events, caused by synoptic weather conditions such as cyclones and frontal system at larger scale, are selected on an hourly basis, by comparing hourly mean values from the gridded reference datasets with defined minimum wind speeds (Table 2, v for strong wind speed cases).

3.2 Statistical evaluation methods

In order to account for spatial differences and displacements between the model and the reference data and to analyze wind speed and direction in a combined way, we apply a novel wind verification methodology. This methodology extends the Fractions Skill Score (FSS), a spatial verification metric developed by Roberts (2008), which is classified as neighborhood-

5 based approach and originally used for verifying precipitation. The FSS is based on the assumption that a model is useful when the model data and the corresponding reference data show a similar spatial frequency of precipitation events, which alliviates the requirement of the models to predict the events at exactly correct positions, which is an unduly strict assumption. Furthermore, this metric avoids the 'double penalty' problem and provides a scale dependent information about the level of model skill (Gilleland et al., 2010; Roberts, 2008).

10 A 'double penalty' problem arises when using traditional statistical methods for datasets which contain an offset between the modeled and the reference data. In that case, the modeled data are penalized twice: first, for simulating an event where it did not occur and second, for failing to simulate an event where it did actually occur (Roberts, 2008; Prein et al., 2013a; Skok and Hladnik, 2018). In order to obtain information of how the skill of a model varies with the spatial scale, the FSS is calculated for different neighborhood sizes. In terms of FSS value, it is a 0-to-1 normalized metric, i.e., the lower limit of the FSS is 0 and

15 the upper limit +1 with values approaching +1 representing a better degree of model performance.

The extended version of the FSS is named the Wind Fractions Skill Sore, denoted as WFSS which has been developed by Skok and Hladnik (2018). The score is calculated based on user defined wind classes. The definition of these classes is partly subjective and can significantly affect the WFSS. The wind vector field should be classified in such a way, that the definition reflects what a user wants to analyze. For example, a complex terrain leads to strong changes in wind directions, therefore it

20 is reasonable to define smaller class intervals regarding the wind directions. For upper level winds the focus could be more on the magnitude of wind speed.

We defined eight wind direction classes with an interval size of $45°$ for a range of three wind speed categories, as shown in the windroses of Fig. 2. Wind speeds $< 0.5\ m\ s^{-1}$ were classified *calm*, independently of the wind direction. The small interval size of $45°$ was chosen to be able to capture the varying wind directions in the study areas, especially for thermally

25 induced winds. Because of the generally much lower wind speeds in WegenerNet FBR we defined a smaller interval size of the wind speed categories for this region (Fig. 2a and b, lower panels) than for the WegenerNet JBT (Fig. 2c and d, lower panels).

Table 3 includes the equations for the calculation of the WFSS and the asymptotic WFSS (AWFSS). This AWFSS will always be reached for a neighborhood size $\geq 2N + 1$, where N is the number of grid points of the largest domain size. At such a large neighborhood size, the estimated fractions within the domain are the same at all locations and further enlarging

30 of this size will not affect the WFSS. A bias always leads to a AWFS value < 1, which indicates systematic differences in the frequency of wind classes between the model and the reference wind classification.

The WFSS is calculated for each hourly field for the selected events. The final case-averaged score values are then estimated based on averaging these hourly WFSS values over the full ensemble of events within the analyzed multi-year period for each evaluation case listed in Table 1.

As briefly explained above, the chosen thresholds of the classification and the number of classes are influencing the score. We found that especially the wind direction thresholds can have a strong impact on the score values. For example, a small change in the wind direction value from prevailing northwesterly winds, which are close to a threshold to distinguish between W and N, could dramatically change the WFSS value. Such a small error in the model data could indicate a poor modeling

5   performance, whereas a human analyst would asses the forecast as reasonably good. To avoid this problem we calculated the hourly WFSS for every rotation between 0° and 45° with an interval size of 5° (nine trail classes), in addition to the original class definition. As next step, always the maximum values of the hourly score values at each neighborhood size are used for computing the final case-averaged score values. We applied this approach for in total 7597 selected events (Table 2, number of events) and estimated the case-averaged score value for each of the eight evaluation cases. A more detailed description of the

10 wind fractions skill score metric can be found in (Skok and Hladnik, 2018).

Furthermore, we applied traditional gridpoint-based statistical performance measures such as bias, root-mean-square error and others to each evaluation case. All statistical performance metrics used in this study are summarized in Table 3.

4   Results

4.1   Evaluation for selected wind events

15 Figure 2a-d illustrates typical examples of modeled wind fields (upper rows of panels) and the corresponding windroses of relative frequency of wind directions divided by wind speed categories (lower rows of panels) from selected representative evaluation events. Each panel depicts modeled and the associated reference data for the WegenerNet FBR (Fig. 2a and b) and the WegenerNet JBT (Fig. 2c and d), for thermally induced (Fig. 2a and c) and strong wind events (Fig. 2b and d). Figure 2e shows the WFSS values of these examples, estimated as explained in Section 3.2 above.

20 The thermally induced wind event on the 29th of July 2009 from 16:00-17:00 UTC for the WegenerNet FBR (Fig. 2a) shows thermally driven regional flows caused by the Alpine pumping. This flow is called Antirandgebirgswind which arises usually in the afternoon as southerly wind with maximum wind speeds of about $2.5\ m\ s^{-1}$ (Wakonigg, 1978). The INCA (upper-left panel of Fig. 2a) and the WegenerNet wind fields (upper-right panel of Fig. 2a) show a similar distribution with generally low wind speeds and prevailing southerly wind directions. The intercomparison of these INCA with WegenerNet data for this

25 event shows the largest WFSS values for all neighborhood sizes, which indicates a good overlap of the wind classes (Fig. 2e, INCAvsWN_therm_FBR). Furthermore, it shows a nearly perfect asymptotic value of about 0.99. This large AWFSS indicates a very small bias, which is also reflected by the similar wind classification results (lower-left and lower-right panel of Fig. 2a).

The COSMO wind field shows similarly low wind speeds, compared to the WegenerNet wind field, but a shift in wind directions from the S sector mainly to the SE and partly to E and NE sectors between the COSMO and the WegenerNet data can

30 be observed (lower-middle and lower-right panel of Fig. 2a).This shift is reflected by small WFSS values at all neighborhood sizes, especially below a scale of 10 km. The AWFSS shows the largest value of all COSMO-CLM intercomparisons, but is still low compared to INCA evaluation cases which indicates a large bias. Evidently, this dynamical wind field modeling at 3-km gridsapcing is not adequately representative for the given challenging hilly terrain.

The strong wind speed event in the WegenerNet FBR on the 30th of October 2008 from 10:00-11:00 UTC led to southwesterly to southerly winds (Fig. 2b). The INCA model data and the WegenerNet reference data show maximum 1-hour vector-mean wind speeds of around $9 - 10 \, m \, s^{-1}$ (upper-left and right panel of Fig. 2b). Regarding the wind directions, differences in the wind sectors can be observed (lower-left and right panel of Fig. 2b). The INCA data show wind directions mainly from the SW

5 sector (lower-left panel) while in the WegenerNet data show wind directions from the S and SW sectors (lower-right panel). The WFSS for this case shows small values at small neighborhood sizes and increases with increasing neighborhood size (Fig. 2e, INCAvsWN_strong_FBR). Despite low WFSS values at small neighborhood sizes caused by differences in wind sectors, the AWFSS shows a high asymptotic value (AWFSS>0.97). This high value is caused by the prevailing wind directions in the WegenerNet data, which are close to the threshold values to distinguish between S and the SW. In this case, the 5° azimuthal

10 class rotation procedure hence avoids lower score values.

Regarding the COSMO data (lower-middle panel of Fig. 2b), no wind is assigned to a wind class with wind speeds higher than $6 \, m \, s^{-1}$, whereas for the WegenerNet wind fields, a large proportion is assigned to the highest wind speed class of this region (lower-right panel of Fig. 2b). This discrepancy leads to the smallest WFSS values at all neighborhood sizes for this region (Fig. 2e, *COSMOvsWN_strong_FBR*) and indicates that the dynamically modeled COSMO wind speeds are

15 systematically underestimated relative to the empirically diagnosed wind speeds.

On the 1th of August 2012 the winds were thermally driven and the local pressure and temperature gradients were causing varying wind speeds and wind directions in the WegenerNet JBT. This is illustrated for the late afternoon INCA and WegenerNet wind fields in the upper left panels of Fig. 2c. The WFSS for the evaluation of the INCA wind field shows the second largest value at the 1 km neighborhood size, which indicates overlapping areas at this neighborhood size, equal to the hor-

20 izontal resolution of the INCA analysis (Fig. 2e, INCAvsWN_therm_JBT). The large AWFSS value indicates again a small bias, which is also reflected by the similar wind classification shown in the windroses of the corresponding lower left panels of Fig. 2c. The high asymptotic value (AWFSS>0.9) indicates a small bias and that the WFSS is mostly influenced by the spatial displacement.

In a further example of a thermally induced wind event on the 31th of May 2008, we intercompare the COSMO with INCA

25 wind fields (right panels of Fig. 2c). In that case, the largest part of the COSMO-CLM modeled flow is from the N and NE sectors, while the INCA system estimated wind directions mainly from the E sector and partly from the NE and SE sectors (bottom right panels of Fig. 2c). This shift in wind directions leads to low WFSS values for all neighborhood sizes including the lowest asymptotic value of all examples, indicating a very poor representation of the wind field by the dynamical modeling in this challenging mountainous terrain.

30 The strong wind speed event for the WegenerNet JBT on the 7th of December 2013 is caused by northwesterly weather conditions. These synoptic scale flow conditions led to strong wind speeds with maximum 1-hour mean wind speeds of around $20 \, m \, s^{-1}$ from 17:00-18:00 UTC. Both the INCA and the WegenerNet wind fields show wind directions mainly from the NW, with some proportions from the N and the W sectors, caused by a channeling of the air flow through the pronounced valleys of this study area. The INCA wind fields show much lower wind speeds in the valley regions compared to the WegenerNet

35 wind fields, resulting from the observations of the ZAMG ADM station that flow into the INCA analysis but are considered

far off the area and not used by the diagnostic modeling (upper left panels of Fig. 2d). As the neighborhood size increases, the WFSS also increases, but due to spatial displacements, the values are generally low (Fig. 2e, INCAvsWN_strong_JBT). The low AWFSS value is caused by the differences in wind speed categories (lower left panels of Fig. 2d).

For the 5th of November 2008 we intercompare the COSMO-CLM wind fields with INCA wind fields from 01:00-02:00 UTC,
5   for a strong wind speed event (right panels of Fig. 2d). In this example, the influence of the smoothed terrain caused by the course horizontal resolution of the COSMO-CLM model becomes obvious. This smoothed topography results in systematically lower wind speeds compared to the INCA wind fields. The WFSS shows similar results like for the previous INCAvsWN_strong_JBT evaluation, with small values at all neighborhood sizes (Fig. 2e, *COSMOvsINCA_strong_JBT*), indicating the clear limits of the COSMO dynamical modeling fields also for strong wind events.

10   **4.2   Statistical evaluation results**

The statistical event-averaged WFSS values from the large ensemble of events over multiple years are represented for each evaluation case in Fig. 3. Overall, it shows a monotonic increase with neighborhood size for all cases so that the AWFSS is the largest value, indication relatively the best performance at large scales.

For the WegenerNet FBR, the statistical WFSS values, calculated for the INCA wind fields compared to the Wegen-
15   erNet wind fields, shows for both the thermally induced and strong wind events nearly the same behavior (Fig. 3a, IN-CAvsWN_therm_FBR and INCAvsWN_strong_FBR). The WFSS values for these cases indicate a reasonably good spatial matching at all neighborhood sizes. Furthermore, the AWFSS values are higher than 0.8, reflecting generally small INCA biases of wind classes.

The statistical WFSS estimated for the COSMOvsWN_therm_FBR case indicates that the COSMO-CLM model clearly
20   and systematically underperforms in case of thermally induced wind events for the WegenerNet FBR. Evidently, due to the coarse horizontal resolution of the wind fields, the COSMO-CLM wind fields appear fundamentally able to capture the varying wind directions for such events in this region. For the COSMOvsWN_strong_FBR case, however the results indicate a similar spatial matching as for the INCAvsWN_therm_FBR and INCAvsWN_strong_FBR cases, just with a somewhat higher bias of wind class differences. This similar performance despite the coarser horizontal resolution of the COSMO model is explained
25   through a weaker influence of the terrain on the wind fields under strong wind conditions in this region.

Because of the challenging terrain of the WegenerNet JBT, the statistical WFSS values are generally low for this region, signalling large biases (Fig. 3b). These biases are indicated by low asymptotic values, which tend to be between 0.61 and 0.64, except for the INCAvsWN_strong_JBT case, which shows an even lower value (AWFSS=0.39).

The spatial displacement and the biases for the INCAvsWN_therm_JBT case are mainly caused by the differences in wind
30   directions for these thermally induced wind events. Especially at small neighborhood sizes at the 1-km scale, WFSS values indicate large spatial displacements.

The INCAvsWN_strong_JBT case shows the lowest values at all neighborhood sizes, but this time caused by the differences in the wind speed categories. These low values are caused by the INCA-analyzed wind speeds, which, in case of strong winds,

are overestimated in the summit regions and underestimated in the valley regions. Slightly overestimated WegenerNet wind speeds in the Enns valley are somewhat reinforcing the difference between the INCA and the WegenerNet wind speeds.

The intercomparison of the COSMO-CLM wind fields with INCA delivers nearly the same (low) WFSS values for both type of wind events. In case of thermally induced events (COSMOvsINCA_therm_JBT) the spatial displacements and biases

5   are mainly caused due to differences in wind directions. For strong wind events (COSMOvsINCA_strong_JBT), the smoothed terrain caused by the coarse resolution of the COSMO-CLM model leads to systematically underestimated wind speeds.

Table 4 summarizes, in addition to the AWFSS values, the results estimated with traditional statistical methods of the INCA analysis and COSMO dynamical fields. Due to the less challenging region of the WegenerNet FBR, all traditional statistical parameters show better performance for this region compared to the WegenerNet JBT. The absolute-value statistical metrics

10   (bias B, standard deviation SD_o, root-mean-square-error $RMSE$) applied to the hourly vector-mean wind speeds, show higher values for the WegenerNet JBT, resulting from the generally higher wind speeds in addition to effects of the complex moun- tainous terrain on the wind fields in this region. The B values are slightly positive for the WegenerNet FBR and negative for the WegenerNet JBT. The substantially negative B value for the INCAvsWN_strong_JBT case again reflects the underestimation of wind speed in the valleys, as explained above. Furthermore, the COSMOvsINCA_strong_JBT intercomparison also shows

15   a negative bias, caused by the coarse resolution of the COSMO-CLM model, which leads to lower wind speeds for strong wind events.

The $RMSE$ values range from 0.79 to 1.85 m s^{-1} for the WegenerNet FBR and from 1.3 to 8.6 m s^{-1} for the WegenerNet JBT. The high value of 8.6 m s^{-1} for the INCAvsWN_strong_JBT case is caused by the underestimation of wind speed in the valleys as well as the overestimation in the summit regions by the INCA model. The mean R values show a better correlation

20   for the WegenerNet FBR than for the WegenerNet JBT. The mean absolute error of wind direction (MAE_{dir}) applied to hourly vector-mean wind directions also shows better performance (INCA and COSMO fields) for the WegenerNet FBR. Due to the varying wind directions caused by thermally induced circulations, the MAE_{dir} is higher for such events for both study areas, with the highest value of 68° for the INCAvsWN_therm_JBT case.

Figure 4 shows the mean wind speed bias spatial distributions for all evaluation cases, for the WegenerNet FBR (Fig. 4a

25   and b) and the WegenerNet JBT (Fig. 4c and d). The distribution for the INCAvsWN_therm_FBR case, the case for thermally induced wind events for the WegenerNet FBR, shows large areas with nearly no B values (left panel of Figure 4a). The maximum B value for this case can be observed in the area of the *Gleichenberger Kogel*, north of the ZAMG Bad Gleichenberg station with a value of around 1.4 m s^{-1}. The evaluation of the COSMO-CLM model for the same thermally induced events shows an overestimation of wind speeds for the whole study area, with B values from 0.5 to 1 m s^{-1} in the western part, and

30   from 1 to about 1.4 m s^{-1} in the eastern part of the study area (right panel of Figure 4a).

The evaluation of the INCA model for strong wind speeds illustrates the strong influence of the terrain on this model. The results show a good agreement in the valleys of the study area, with partly small negative B values (left panel of Figure 4b). The hilltop regions exhibit positive B values, with maximum values of around 5 m s^{-1} again in the area of the *Gleichenberger Kogel*. Overall positive B values of COSMO-CLM dynamical wind speed fields for strong wind events are seen in the right panel of

35   Figure 4b, showing the systematic overestimating by COSMO fields in this case.

For the WegenerNet JBT, the strong influence of the terrain on the INCA-analyzed wind speeds can be observed in all evaluation cases in this region. The evaluation of the INCA model for thermally induced wind events exhibit negative B values in the valleys, whereby positive values are partly occurring in the summit regions (left panel of Figure 4c). At lower elevations, the intercomparison of the COSMO-CLM with the INCA model shows nearly no B values for thermally induced events (right panel of Figure 4c). Furthermore, small negative B values at the summit regions and some spots with positive values can be observed for this case.

Similar bias distribution patterns as for the INCA evaluation for thermally induced wind events are present for strong wind events in the INCAvsWN_strong_JBT case, but this time with strong negative and positive B values ranging from -14.4 to $4.9\ m\ s^{-1}$ (left panel of Figure 4d). These strong negative B values are again caused by the severely underestimated INCA wind speeds and the somewhat overestimated WegenerNet wind speeds in the valley regions of the WegenerNet JBT.

Opposite patterns can be seen for the intercomparison of the COSMO-CLM with the INCA model. This intercomparison exhibits small positive values in the valley regions and strong negative values in the summit regions (right panel of Figure 4d). Main reason for these strong negative B values is the course resolution of the COSMO-CLM data and the resulting underestimation of wind speeds for strong wind events, as explained above.

5 Conclusions

In this work we evaluated wind fields generated by two different modeling systems against empirically diagnosed wind fields from WegenerNet high-density network data: the INCA analysis system of the Austrian weather service ZAMG (Haiden et al., 2011) and the non-hydrostatic COSMO-CLM model (Schättler et al., 2016). The INCA wind fields have a horizontal resolution of 1 km × 1 km, and in case of COSMO-CLM, 3 km × 3 km horizontal resolution was available, both on an hourly basis. The empirical high-resolution wind fields from the WegenerNet where generated by the WegenerNet Wind Product Generator (WPG), recently developed by Schlager et al. (2017, 2018).

The WPG-diagnosed gridded wind fields are available with a temporal resolution of 30 minutes and a spatial resolution of 100 m × 100 m and can therefore well serve as reference. The WegenerNet Feldbach Region (FBR) was used as study area, characterized by generally small differences in altitude in the hilly terrain of this region. The second study area was the WegenerNet Johnsbachtal (JBT) region, which is a mountainous region characterized through a very complex terrain.

The evaluation of the INCA and the COSMO-CLM wind fields was based on classifying the data separately into thermally induced and strong wind events. In case of the INCA evaluation, we could select wind events within the period 2008-2017 for the WegenerNet FBR and within 2012-2017 for the WegenerNet JBT. For evaluating the COSMO-CLM data, events from the period 2008-2010 were selected for both study areas. Due to WegenerNet JBT wind fields being not yet available within 2008-2010, we intercompared the COSMO-CLM wind fields with INCA wind fields in this region.

Besides traditional performance measures such as bias, root-mean-square error, correlation coefficient, and mean absolute error of wind direction, we in particular applied a spatial wind verification methodology named the Wind Fractions Skill

Score (WFSS) (Skok and Hladnik, 2018). This new score was used to detect spatial displacements of wind patterns and biases based on predefined wind speed and direction classes. The WFSS avoids the 'double penalty' problem and is able to distinguish between a 'near miss' and large displacements between modeled and reference wind fields. Furthermore, a spatial scale-dependent skill is determined by this score.

5   Due to the less challenging terrain of the Alpine foreland region, all statistical performance measures showed better INCA and COSMO performance for the WegenerNet FBR than for the challenging mountainous region WegenerNet JBT. The spatial verification of all evaluation cases indicates an increasing skill with increasing by larger scale (neighborhood size). For both study areas, the traditional statistical performance measures, applied to the wind speed, mostly show better performance of INCA and COSMO for thermally induced wind events than for strong wind events. On the other hand, the results related to
10   wind direction indicate a better performance for strong wind events than for thermally induced events.

     More specifically, the verification for the WegenerNet FBR shows that the INCA analysis wind fields are more skillful than the COSMO-CLM dynamical wind fields in this region. The INCA verification indicates a reasonably good performance for both thermally induced and strong wind events, with some spatial displacements at smaller scales and small biases in wind classes. Regarding wind speed, we found a generally good modeling performance by these analysis fields for both types of
15   wind events with slight overestimates in the summit regions of the WegenerNet FBR.

     The COSMO-CLM dynamical model clearly performs less well in case of thermally induced wind events for this region. The reason for this weak performance is the limited resolution of the wind field dataset from this model. With the resolution of 3 km × 3 km, the fields are not able to resolve the varying wind speeds and directions caused by thermally driven circulations. The wind speeds are overestimated by this model for both, thermally and strong events, and large differences in wind directions
20   are found for thermally induced events.

     For the WegenerNet JBT region, the verification shows generally large spatial displacements at all scales and strong biases in wind classes. In case of the INCA evaluation, large wind direction deviations for thermally induced wind events indicate that the analysis fields are not able to adequately capture the varying wind patterns such as slope and valley winds, which roots in the sparse station density that INCA can anchor to and the coarse horizontal resolution of the first guess provided by the
25   ALARO model in this complex-terrain region. Furthermore, the statistics show an substantial underestimation of wind speeds in the valleys and overestimated wind speeds in parts of the summit regions for both type of wind events. Especially for strong wind events we found a large negative bias in the valley regions and a positive bias in parts of the summit regions. Somewhat overestimated WegenerNet empirical wind speeds in the region of the Enns valley contribute in a minor way to the strong negative bias.

30   The intercomparison of the COSMO-CLM dynamical fields with the INCA analysis fields for thermally induced wind events shows large displacements and a bias from wind speed class differences. The error measure applied to the wind direction for this intercomparison indicates large differences in wind directions for thermally induced wind events between the COSMO and INCA wind fields. Furthermore, large spatial displacements, biases in populating the wind classes, underestimated wind speeds in the summit regions, and slightly overestimated wind speeds in the valleys, are reflecting the disadvantage of smoothed
35   terrain, which is caused by the limited resolution of 3 km × 3 km of the COSMO-CLM model.

Based on these finding we suggest, underpining the results of Haiden et al. (2011), that additional observed wind information in the summit and valley regions, especially in a complex terrain like the WegenerNet JBT, and a more comprehensive use of wind-constraining satellite data as well as a higher-resolution RCM could help to systematically improve the INCA-analyzed wind fields.

Related to the COSMO dynamical modeling, a verification of COSMO-CLM-generated higher-resolution 1 km × 1 km wind fields would also be a promising issue for further investigations of how this may improve the modeling of wind patterns in a complex terrain.

Investigations regarding the WegenerNet JBT wind fields showed, that an additional wind-observing station in the Enns valley would improve the results for this region (Schlager et al., 2018). Such an additional station would avoid the overestimation of WegenerNet wind speeds in the Enns valley, especially for strong wind events. In the WegenerNet FBR region just recently (in May 2018) another wind station was added in the Raab valley (station Nr. 155 1b), which will further improve the WPG-derived fields in future. This adds further value to valuable reference for evaluation of important other data products such as the INCA operational analysis and dynamical climate model fields.

Code availability. The CALMET 6.5.0 model code is available from the website www.src.com/calpuff/. The INCA and the WPG code is not in the public domain and cannot be distributed. The source code for the COSMO-CLM is available on request via the website https://www.clm-community.eu. The code for the calculation of FSSwind score is available as part of the SpatialVx package (function calculate_FSSwind). SpatialVx is a R software package made by Eric Gilleland that enables the calculation of a large number of spatial verification scores (https://cran.r-project.org/package=SpatialVx).

Data availability. CORINE Land Cover data for the study area were taken from www.eea.europa.eu, digital elevation model data from www.gis.steiermark.at, and WegenerNet data from www.wegenernet.org. The WegenerNet data contain the WPG wind field output data as introduced in this study. The INCA data are available on request from the Central Institute for Meteorology and Geodynamics (klima@zamg.ac.at). COSMO climate model data are available on request from the Wegener Center for Climate and Global Change (heimo.truhetz@uni-graz.at).

Author contributions. C. Schlager collected the data, performed the analyses and modeling, created the figures, and wrote the first draft of the manuscript. G. Kirchengast provided guidance and advice on all aspects of the study and significantly contributed to the text. J. Fuchsberger provided guidance on technical aspects of the WegenerNet networks, and its data characteristics and contributed to the text. A. Kann provided INCA-related advice and contributed to the INCA part of the text and H. Truhetz provided information and advice on the COSMO-CLM model setup and characteristics. All authors commented on the final version of the manuscript.

Competing interests. We declare that no competing interests are present.

Acknowledgements. The authors thank Gregor Skok (University of Ljubljana, Department of Physic), for providing the R code to calculate the Wind Fractions Skill Score. Furthermore, the authors acknowledge the data providers at the Central Institute for Meteorology and Geodynamics (ZAMG) for the Integrated Nowcasting through Comprehensive Analysis (INCA) dataset. WegenerNet funding is provided by the Austrian Ministry for Science and Research, the University of Graz, the state of Styria (which also included European Union regional

5   development funds), and the city of Graz; detailed information can be found online (www.wegcenter.at/wegenernet).

References

Abdel-Aal, R., Elhadidy, M., and Shaahid, S.: Modeling and forecasting the mean hourly wind speed time series using GMDH-based abductive networks, Renewable Energy, 34, 1686–1699, https://doi.org/10.1016/j.renene.2009.01.001, 2009.

Awan, N. K., Truhetz, H., and Gobiet, A.: Parameterization-induced error characteristics of MM5 and WRF operated in climate mode over the alpine region: An ensemble-based analysis, J. Climate, 24, 3107–3123, https://doi.org/10.1175/2011JCLI3674.1, 2011.

Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., Vitart, F., and Balsamo, G.: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales, Quart. J. Roy. Meteor. Soc., 134, 1337–1351, https://doi.org/10.1002/qj.289, http://dx.doi.org/10.1002/qj.289, 2008.

Gilleland, E., Ahijevych, D. A., Brown, B. G., and Ebert, E. E.: Verifying Forecasts Spatially, Bull. Amer. Meteor. Soc., 91, 1365–1376, https://doi.org/10.1175/2010BAMS2819.1, 2010.

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, https://www.geosci-model-dev.net/8/3349/2015/, 2015.

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C.: The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the eastern alpine region, Wea. Forecasting, 26, 166–183, https://doi.org/10.1175/2010WAF2222451.1, 2011.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, J. Geophys. Res., 113, n/a–n/a, https://doi.org/10.1029/2008JD010201, 2008.

Hiebl, J. and Frei, C.: Daily temperature grids for Austria since 1961—concept, creation and applicability, Theor. Appl. Climatol., 124, 161–178, https://doi.org/10.1007/s00704-015-1411-4, 2016.

Hohmann, C., Kirchengast, G., and Birk, S.: Alpine foreland running drier? Sensitivity of a drought vulnerable catchment to changes in climate, land use, and water management, Climatic Change, 147, 179–193, https://doi.org/10.1007/s10584-017-2121-y, 2018.

Kabas, T.: WegenerNet climate station network region Feldbach: Experimental setup and high resolution data for weather and climate research (in German), Scientific Rep. 47-2012, Wegener Center Verlag, Graz, Austria, [Available online at http://wegcwww.uni-graz.at/publ/wegcreports/2012/WCV-WissBer-No47-TKabas-Jan2012.pdf.], 2012.

Kabas, T., Foelsche, U., and Kirchengast, G.: Seasonal and annual trends of temperature and precipitation within 1951/1971-2007 in southeastern Styria, Austria, Meteor. Z., 20, 277–289, https://doi.org/10.1127/0941-2948/2011/0233, 2011.

Kann, A., Meirold-Mautner, I., Schmid, F., Kirchengast, G., Fuchsberger, J., Meyer, V., Tuechler, L., and Bica, B.: Evaluation of high-resolution precipitation analyses using a dense station network, Hydrol. Earth Syst. Sci., 19, 1547–1559, https://doi.org/10.5194/hess-19-1547-2015, 2015a.

Kann, A., Wittmann, C., Bica, B., and Wastl, C.: On the Impact of NWP Model Background on Very High–Resolution Analyses in Complex Terrain, Wea. Forecasting, 30, 1077–1089, https://doi.org/10.1175/WAF-D-15-0001.1, 2015b.

Kendon, E. J., Ban, N., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., Evans, J. P., Fosser, G., and Wilkinson, J. M.: Do convection-permitting regional climate models improve projections of future precipitation change?, Bull. Amer. Meteor. Soc., 98, 79–93, https://doi.org/10.1175/BAMS-D-15-0004.1, 2017.

Kirchengast, G., Kabas, T., Leuprecht, A., Bichler, C., and Truhetz, H.: WegenerNet: A pioneering high-resolution network for monitoring weather and climate, Bull. Amer. Meteor. Soc., 95, 227–242, https://doi.org/10.1175/BAMS-D-11-00161.1, 2014.

Leutwyler, D., Fuhrer, O., Lapillonne, X., Lüthi, D., and Schär, C.: Towards European-scale convection-resolving climate simulations with GPUs: A study with COSMO 4.19, Geosci. Model Dev., 9, 3393–3412, https://doi.org/10.5194/gmd-9-3393-2016, 2016.

Lugauer, M. and Winkler, P.: Thermal circulation in South Bavaria – climatology and synoptic aspects, Meteor. Z., 14, 15–30, https://doi.org/10.1127/0941-2948/2005/0014-0015, 2005.

5 O, S., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., and Petersen, W. A.: Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria, Hydrol. Earth Syst. Sci., 21, 6559–6572, https://doi.org/10.5194/hess-21-6559-2017, 2017.

O, S., Foelsche, U., Kirchengast, G., and Fuchsberger, J.: Validation and correction of rainfall data from the WegenerNet high density network in southeast Austria, J. Hydrol., 556, 1110–1122, https://doi.org/10.1016/j.jhydrol.2016.11.049, 2018.

10 Osborn, T. J. and Hulme, M.: Evaluation of the European daily precipitation characteristics from the atmospheric model intercomparison project, Int. J. Climatol., 18, 505–522, https://doi.org/10.1002/(SICI)1097-0088(199804)18:5<505::AID-JOC263>3.0.CO;2-7, 1998.

Prein, A. F.: Added Value of Convection Resolving Climate Simulations, Scientific Report 53-2013, Wegener Center Verlag, Graz, Austria, [Available online at http://wegcwww.uni-graz.at/publ/wegcreports/2013/WCV-SciRep-No53-APrein-Jul2013.pdf.], 2013.

Prein, A. F., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N. K., Keuler, K., and Georgievski, G.: Added value of convection
15 permitting seasonal simulations, Climate Dyn., 41, 2655–2677, https://doi.org/10.1007/s00382-013-1744-6, https://doi.org/10.1007/s00382-013-1744-6, 2013a.

Prein, A. F., Holland, G. J., Rasmussen, R. M., Done, J., Ikeda, K., Clark, M. P., and Liu, C. H.: Importance of regional climate model grid spacing for the simulation of heavy precipitation in the colorado headwaters, J. Climate, 26, 4848–4857, https://doi.org/10.1175/JCLI-D-12-00727.1, 2013b.

20 Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Toelle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., Van Lipzig, N. P., and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Rev. Geophys., 53, 323–361, https://doi.org/10.1002/2014RG000475, 2015.

Prettenthaler, F., Podesser, A., and Pilger, H.: Climate Atlas Styria, Period 1971-2000: An Application-Oriented Climatology (in German), vol. 4, Verlag der Oesterreichischen Akademie der Wissenschaften, Wien, 2010.

25 Roberts, N.: Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model, Meteor. Appl., 15, 163–169, https://doi.org/10.1002/met.57, http://dx.doi.org/10.1002/met.57, 2008.

Rockel, B., Will, A., and Hense, A.: The Regional Climate Model COSMO-CLM (CCLM), Meteor. Z., 17, 347–348, https://doi.org/10.1127/0941-2948/2008/0309, http://dx.doi.org/10.1127/0941-2948/2008/0309, 2008.

Schättler, Doms, G, U., and Baldauf, M.: A Description of the Nonhydrostatic Regional COSMO Model; Part VII: User's Guide, Deutscher
30 Wetterdienst, 3004 Offenbach, Germany, 2016.

Schlager, C., Kirchengast, G., and Fuchsberger, J.: Generation of high-resolution wind fields from the dense meteorological station network WegenerNet in south-eastern Austria, Wea. Forecasting, 32, 1301–1319, https://doi.org/10.1175/WAF-D-16-0169.1, 2017.

Schlager, C., Kirchengast, G., and Fuchsberger, J.: Empirical high-resolution wind field and gust model in mountainous and hilly terrain based on the dense WegenerNet station networks, Atmos. Meas. Tech., 2018, 1–32, https://doi.org/10.5194/amt-2018-31, https://www.
35 atmos-meas-tech-discuss.net/amt-2018-31/, in press, 2018.

Schroeer, K. and Kirchengast, G.: Sensitivity of extreme precipitation to temperature: the variability of scaling factors from a regional to local perspective, Climate Dyn., 50, 3981–3994, 2018.

Sfetsos, A.: A novel approach for the forecasting of mean hourly wind speed time series, Renewable Energy, 27, 163 – 174, https://doi.org/http://dx.doi.org/10.1016/S0960-1481(01)00193-8, 2002.

Skok, G. and Hladnik, V.: Verification of Gridded Wind Forecasts in Complex Alpine Terrain: A New Wind Verification Methodology Based on the Neighborhood Approach, Mon. Wea. Rev., 146, 63–75, https://doi.org/10.1175/MWR-D-16-0471.1, 2018.

5 Strasser, U., Marke, T., Sass, O., Birk, S., and Winkler, G.: John's creek valley: A mountainous catchment for long-term interdisciplinary human-environment system research in Upper Styria (Austria), Environ. Earth Sci., 69, 695–705, https://doi.org/10.1007/s12665-013-2318-y, 2013.

Suklitsch, M., Gobiet, A., Truhetz, H., Awan, N. K., Göttel, H., and Jacob, D.: Error characteristics of high resolution regional climate models over the Alpine area, Climate Dyn., 37, 377–390, https://doi.org/10.1007/s00382-010-0848-5, 2011.

10 Wakonigg, H.: Weather and Climate in Styria (in German), Verlag fuer die Technische Universitaet Graz, Graz, 1978.

Wang, Y., Haiden, T., and Kann, A.: The operational Limited Area Modelling system at ZAMG: ALADIN-AUSTRIA, vol. 37, Österreichische Beiträge zu Meteorologie und Geophysik, Wien, 2006.

Whiteman, C.: Mountain Meteorology: Fundamentals and Applications, Oxford University Press, Graz, 2000.
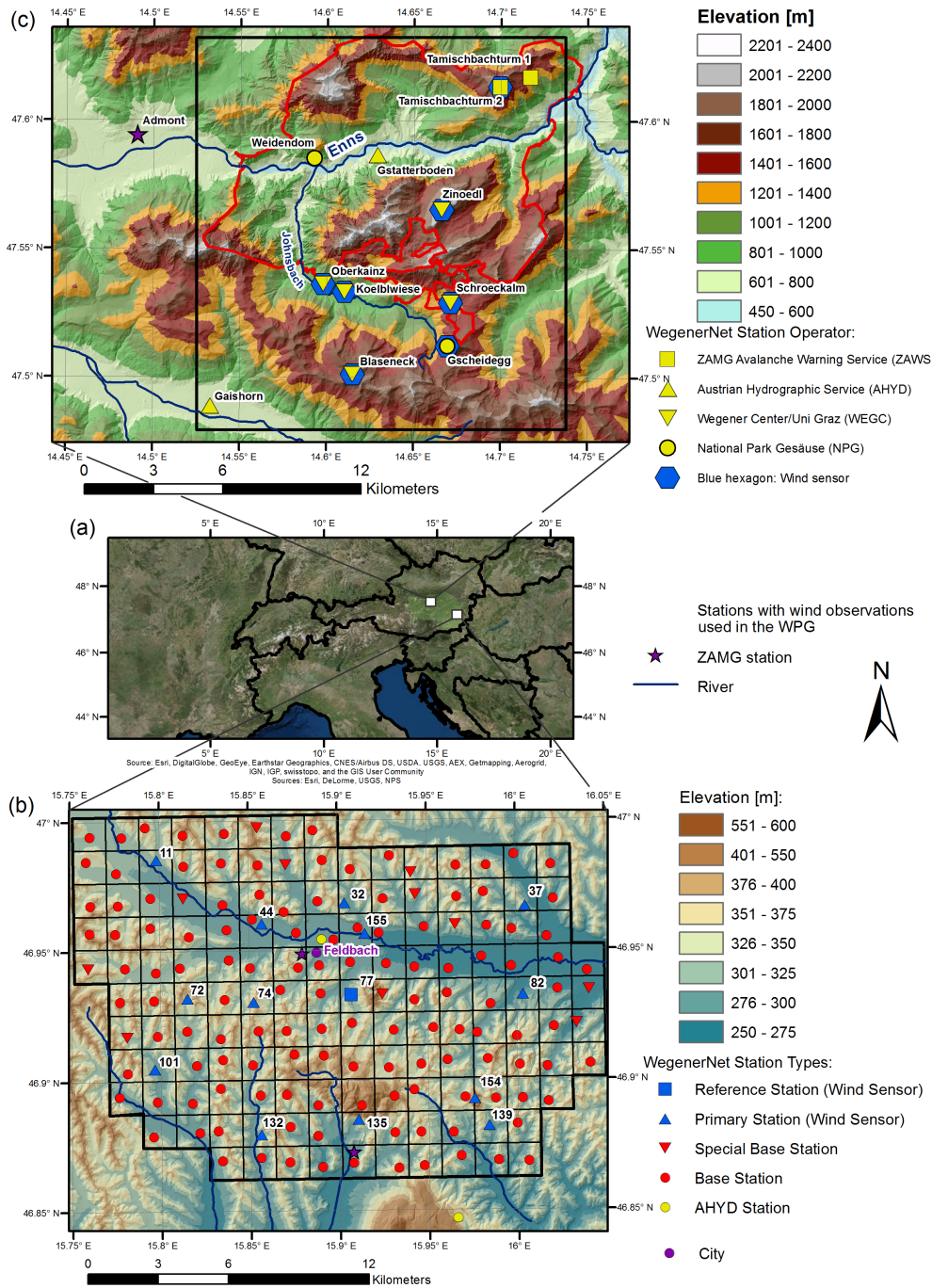
Figure 1. (a) Location of the study areas, the WegenerNet Felbach Region (FBR) in the southeast of Styria, Austria and the WegenerNet Johnsbachtal (JBT) region in the north of Styria, Austria (white-filled rectangles, enlarged in (b) and (c)). (b) The WegenerNet FBR with its 154 meteorological stations, with the legend explaining map characteristics and station types. (c) Map of the WegenerNet JBT region (black rectangle) including its meteorological stations, with the legend explaining map characteristics and station operators.
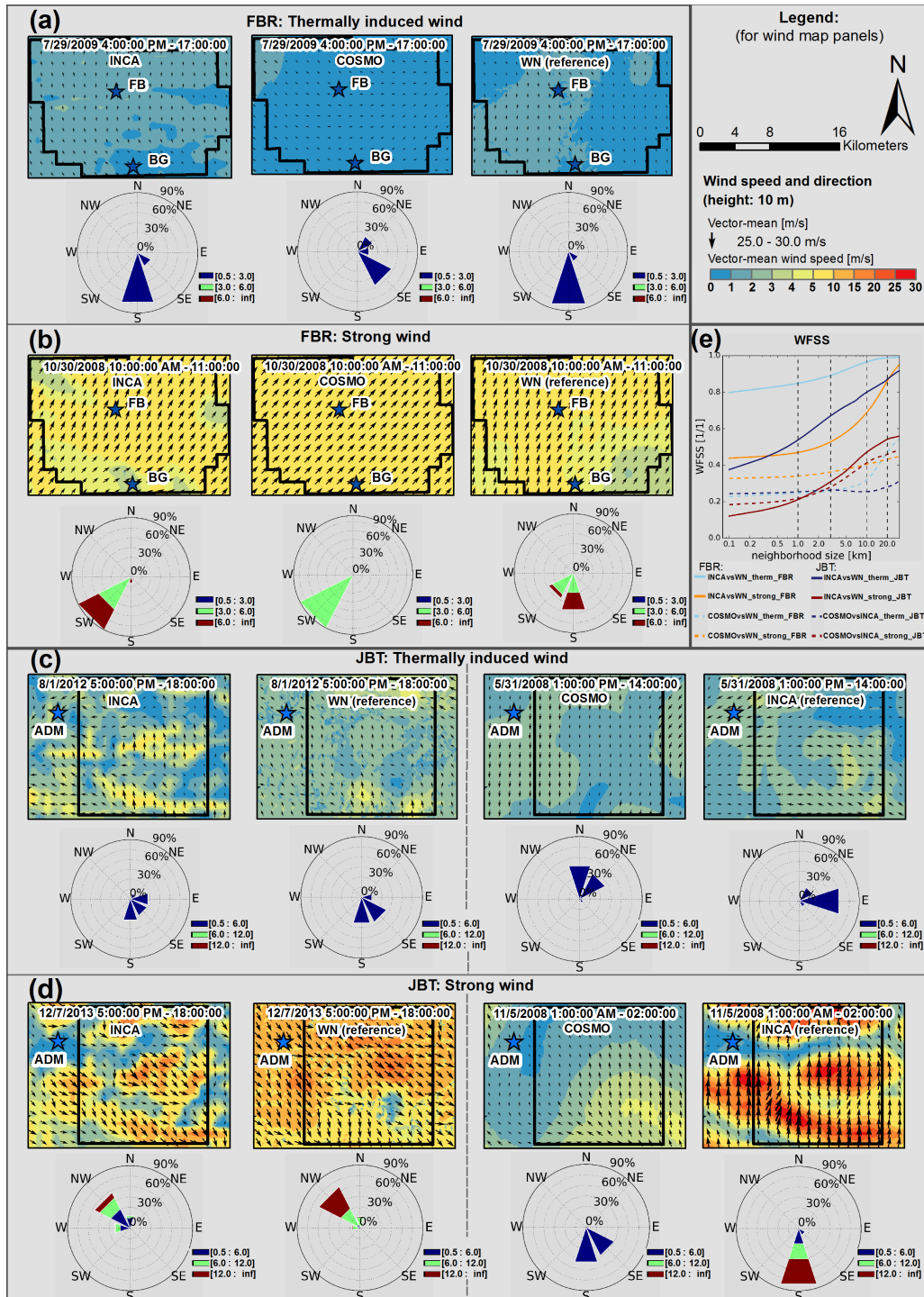
Figure 2. Wind Fractions Skill Score (WFSS) analysis for selected one-hour wind fields for the WegenerNet FBR (a, b) and the WegenerNet JBT region (c, d). (a-d) Modeled and reference wind fields (first row) and corresponding relative frequency of wind directions for a range of wind speed categories (second row), in each panel. (e) WFSS results for the modeled versus reference wind fields from (a)-(d). See Table 1 for more information on the evaluation cases.
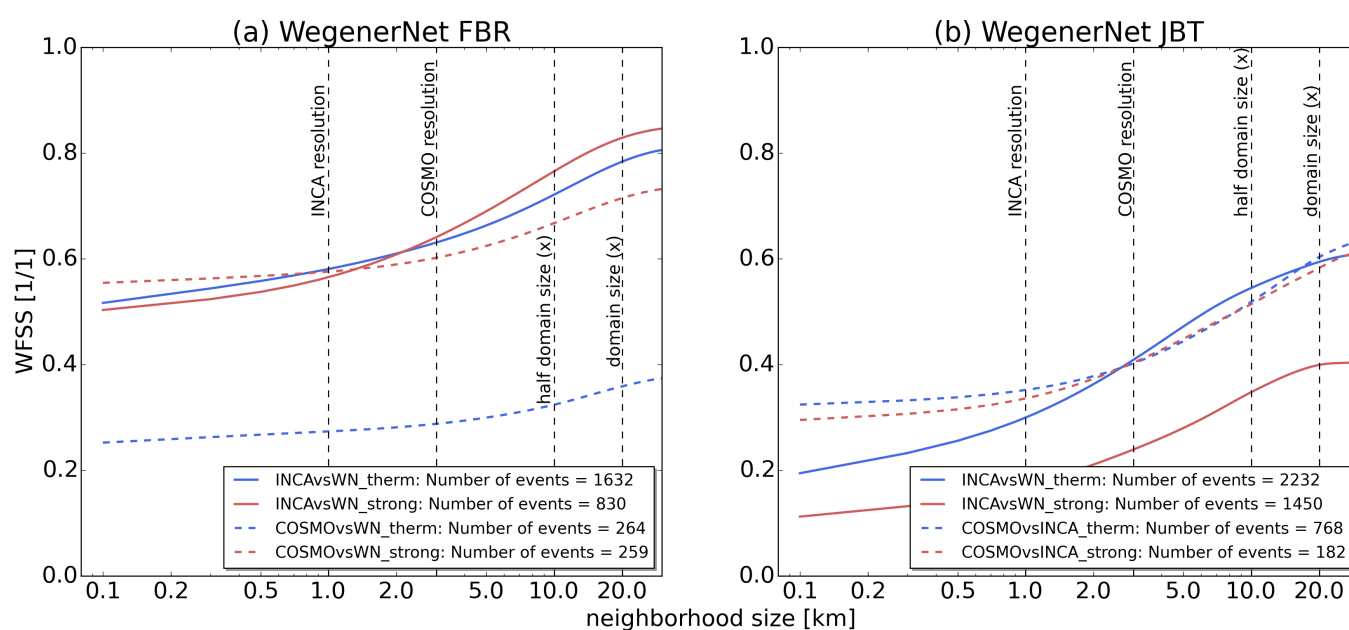
Figure 3. The event averaged Wind Fractions Skill Score (WFSS) results for the WegenerNet FBR (a), compared to the WegenerNet JBT (b), for the four defined evaluation cases in each region (see legend; indicating also the number of events included). See Table 1 for more information on the evaluation cases.
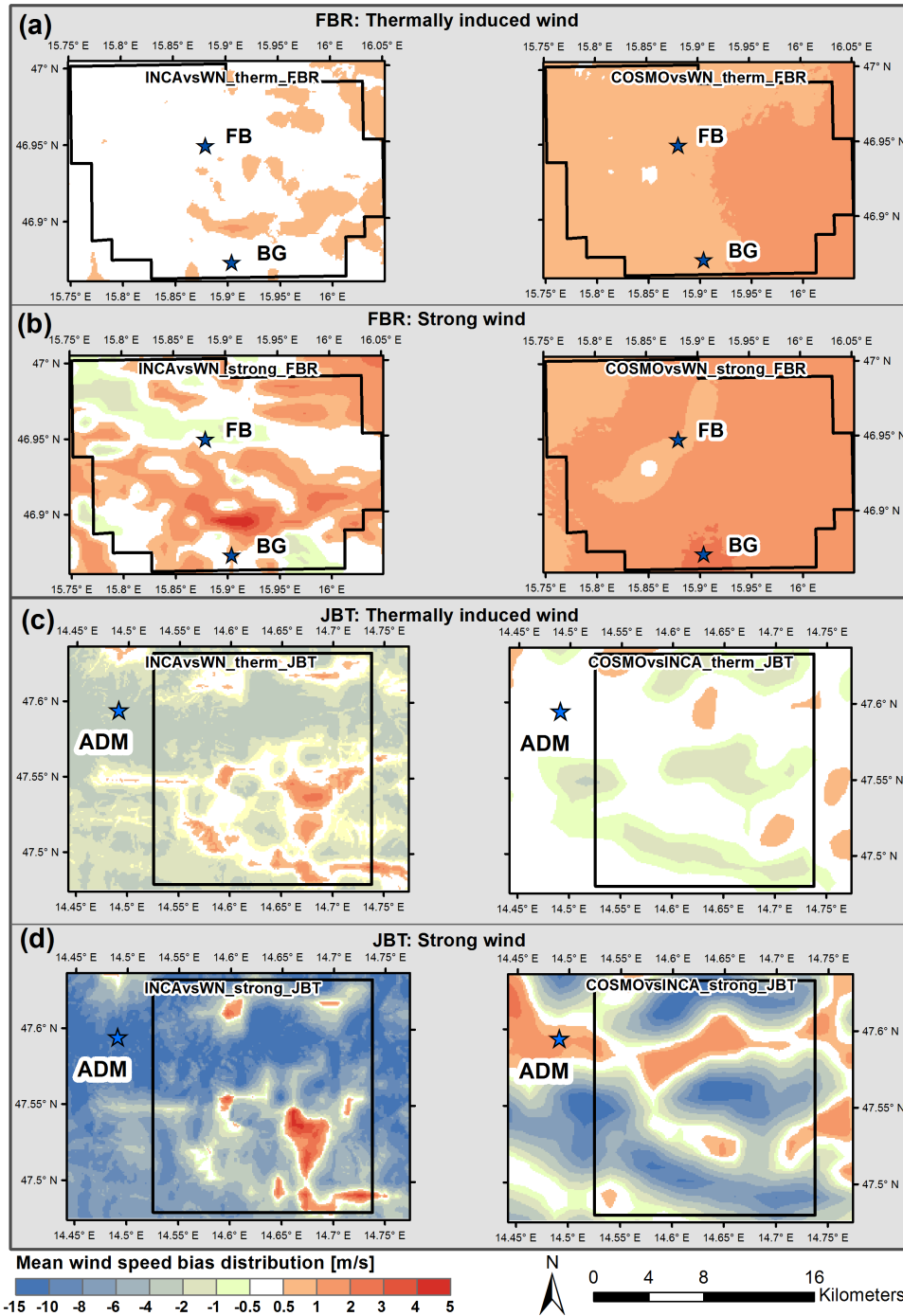
Figure 4. Mean wind speed bias distribution for the WegenertNet FBR (a, b) and the WegenerNet JBT (c, d): (a, b) INCA versus WegnerNet (left) and COSMO versus WegenerNet (right) for (a) thermally induced wind events and (b) strong wind events, and (c, d) WegenerNet versus INCA (left) and COSMO versus INCA (right) for (c) thermally induced wind events and (d) strong wind events.

Table 1. Characteristics of wind field evaluation cases used for the WegenerNet, INCA, and COSMO intercomparisons (top half for the WegenerNet FBR; bottom half for the WegenerNet JBT region).

Evaluation Case	Type	Region	Modeled dataset	Reference dataset	Period
WegenerNet FBR					
INCAvsWN_therm_FBR	thermally	FBR	INCA	WN	2008-2017
INCAvsWN_strong_FBR	strong	FBR	INCA	WN	2008-2017
COSMOvsWN_therm_FBR	thermally	FBR	COSMO	WN	2008-2010
COSMOvsWN_strong_FBR	strong	FBR	COSMO	WN	2008-2010
WegenerNet JBT					
INCAvsWN_therm_JBT	thermally	JBT	INCA	WN	2012-2017
INCAvsWN_strong_JBT	strong	JBT	INCA	WN	2012-2017
COSMOvsINCA_therm_JBT	thermally	JBT	COSMO	INCA	2008-2010
COSMOvsINCA_strong_JBT	strong	JBT	COSMO	INCA	2008-2010

Table 2. Limits for the selection of thermally induced or strong wind events for the defined evaluation cases shown in Table 1 (top half for the WegenerNet FBR; bottom half for the WegenerNet JBT).

Evaluation Case	*Variables*[a]	$v\,[\mathrm{m\,s^{-1}}]$ *(reference data*[b]*)*	$rh\,[\%]$ *(reference data*[b]*)*	$Q_{g,m} - Q_{n(g),o}\,[\mathrm{W\,m^{-2}}]$ *(reference data*[b]*)*	$Q_n\,[\mathrm{W\,m^{-2}}]$ *(reference data*[b]*)*	*Number of events*
			WegenerNet FBR			
INCAvsWN_therm_FBR		<1.5 (RS$_{dm}$)	<65.0 (RS$_{dm}$)	<100.0 (RS$_{dm}$)	<30.0 (RS$_{nm}$)	1632
INCAvsWN_strong_FBR		>3.0 (WN$_{hm}$)	–	–	–	830
COSMOvsWN_therm_FBR		<1.5 (RS$_{dm}$)	<65.0 (RS$_{dm}$)	<100.0 (RS$_{dm}$)	<30.0 (RS$_{nm}$)	1632
COSMOvsWN_strong_FBR		>3.0 (WN$_{hm}$)	–	–	–	259
			WegenerNet JBT			
INCAvsWN_therm_JBT		<2.0 (SCH$_{dm}$)	<65.0 (SCH$_{dm}$)	<20.0 (SCH$_{dm}$)	<30.0 (SCH$_{nm}$)	2232
INCAvsWN_strong_JBT		>9.0 (INCA$_{hm}$)	–	–	–	1450
COSMOvsINCA_therm_JBT		–	<65.0 (WEI$_{dm}$)	<20.0 (WEI$_{dm}$)	–	768
COSMOvsINCA_strong_JBT		>6.0 (INCA$_{hm}$)	–	–	–	182

[a]v: wind speed; rh: relative humidity; $Q_{g,m}$ - $Q_{(n)g,o}$: difference between mean modeled global radiation ($Q_{g,m}$) and observed net radiation ($Q_{n,o}$) for the WegenerNet FBR and difference between $Q_{g,m}$ and observed global radiation ($Q_{g,o}$) for the WegenerNet JBT; Q_n: net radiation.

[b]RS: Reference Station (Nr. 77); SCH: Schroeckalm station; WEI: Weidendom station; dm: daytime mean value from observations (from sunrise till sunset); nm: nighttime mean value from observations (from sunset till sunrise); WN$_{hm}$: hourly mean value from gridded WegenerNet wind speed; INCA$_{hm}$: hourly mean value form gridded INCA wind speed.

Table 3. Statistical performance parameters used for the intercomparison of the wind field modeling results.

Parameter	Equation	Remarks
Wind fractions skill score	$WFSS = 1 - \dfrac{\sum_k \sum_{i,j} [O_k(i,j) - M_k(i,j)]^2}{\sum_k \left\{ \sum_{i,j} O_k(i,j)^2 + \sum_{i,j} M_k(i,j)^2 \right\}}$	O_k: fraction values for observations for wind class k at location i,j; M_k: fraction values for forecasts for wind class k at location i,j (Skok and Hladnik, 2018; Roberts, 2008)
Asymptotic WFSS	$AWFSS = 1 - \dfrac{\sum_k \left(f_k^O - f_k^M\right)^2}{\sum_k \left\{ \left(f_k^O\right)^2 + \left(f_k^M\right)^2 \right\}}$	f_k^O: frequency of wind class k in the observations; f_k^M: frequency of wind class k in the forecast (Skok and Hladnik, 2018)
Bias	$B = \frac{1}{N} \sum_{i=1}^N \left(v_{m,i} - v_{o,i}\right)$	v_m: modeled wind speed; v_o: observed wind speed
Standard deviation of observed wind speed	$SD_o = \sqrt{\frac{1}{(N-1)} \sum_{i=1}^N (v_{o,i} - \overline{v_o})^2}$	v_o: observed wind speed; $\overline{v_o}$: mean observed wind speed
Root-mean-square error	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (v_{m,i} - v_{o,i})^2}$	v_m: modeled wind speed; v_o: observed wind speed
Correlation coefficient	$R = \frac{1}{(N-1)} \sum_{i=1}^N \left(\frac{v_{m,i} - \overline{v_m}}{\sigma_m}\right) \left(\frac{v_{o,i} - \overline{v_o}}{\sigma_o}\right)$	v_m: modeled wind speed; $\overline{v_m}$: mean modeled wind speed; v_o: observed wind speed; $\overline{v_o}$: mean observed wind speed; σ_m: standard deviation of modeled wind speed; σ_o: standard deviation of observed wind speed
Mean absolute error of wind direction	$MAE_{\text{dir}} = \frac{1}{N} \sum_{i=1}^N \left\{ \arccos[\cos(\phi_{m,i} - \phi_{o,i})] \right\}$	ϕ_m: modeled wind direction; ϕ_o: observed wind direction

Table 4. Statistical performance measures calculated for the evaluation cases from Table 1, for the WegenerNet FBR and the WegenerNet JBT region. See Table 3 for more information on the calculation of the parameters.

Evaluation Case	$AWFSS$ [1/1]	B [m s^{-1}]	SD_o [m s^{-1}]	$RMSE$ [1/1]	R [1/1]	MAE_{dir} [°]
			WegenerNet FBR			
INCAvsWN_therm_FBR	0.81	0.34	0.74	0.79	0.67	38
INCAvsWN_strong_FBR	0.85	0.50	1.04	1.66	0.34	14
COSMOvsWN_therm_FBR	0.38	1.32	0.72	1.85	0.37	55
COSMOvsWN_strong_FBR	0.74	1.01	1.03	1.28	0.57	14
Mean Value	0.70	0.79	0.88	1.40	0.49	30
			WegenerNet JBT			
INCAvsWN_therm_JBT	0.61	-1.37	2.37	2.97	0.20	68
INCAvsWN_strong_JBT	0.39	-6.69	3.97	8.60	0.16	39
COSMOvsINCA_therm_JBT	0.64	-0.23	1.32	1.31	0.40	56
COSMOvsINCA_strong_JBT	0.63	-3.79	4.24	5.52	0.08	25
Mean Value	0.57	-3.04	2.98	4.62	0.20	47