

RESEARCH ARTICLE

Wind-Topo: Downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning

Jérôme Dujardin^{1,2}  | Michael Lehning^{1,2}

¹School of Architecture, Civil and Environmental Engineering, Swiss Federal Institute of Technology in Lausanne (EPFL), Lausanne, Switzerland

²Institute for Snow and Avalanche Research (SLF), Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Davos, Switzerland

Correspondence

J. Dujardin, Institute for Snow and Avalanche Research (SLF), Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Davos 7260, Switzerland.
Email: jerome.dujardin@slf.ch

Funding information

Innosuisse through the Swiss Competence Center for Energy Research, Supply of Electricity

Abstract

Predicting wind flow in highly complex terrain like the Alps is a challenge for all models. When physical processes need to be resolved in a spatially explicit manner, grids with high horizontal resolution of a few hundred meters are often required and drastically limit, in many cases, the extent and duration of the simulations. Many surface process models, like the simulation of heterogeneous snow cover across a season, however, need long time series on large domains as inputs. Statistical downscaling can provide the required data, but no model can reach the desired resolutions effectively and provide temporally resolved wind speed and direction on highly complex topography. The assessment of the potential for wind energy in the Alps, a promising player in the energy transition, is an example where the current shortcomings cause strong limitations. We present “Wind-Topo”, a novel approach based on deep learning that discovers some of the interactions between high-resolution topography and coarser-resolution states of the atmosphere to generate near-surface wind fields with a 50-m resolution. In our test case, we use a large number of measurement stations in Switzerland to train the model and an operational weather prediction model (COSMO-1) as predictor. Wind-Topo employs a custom architecture that analyses the state of the atmosphere on various scales and associates it with high-resolution topography. A dedicated loss function leads to good scoring metrics as well as accurate wind-speed distributions at 60 independent stations used for a thorough validation. 50-m resolution wind fields are generated efficiently and exhibit several expected orographic effects like ridge acceleration, sheltering, and deflection. Furthermore, the bias and mean absolute error from COSMO-1 at the alpine validation stations, which are 0.72 and 1.77 m·s⁻¹ respectively, are reduced to -0.07 and 1.21 m·s⁻¹.

KEYWORDS

complex terrain, convolutional neural network, deep learning, downscaling, orographic effect, wind

1 | INTRODUCTION

In complex terrain, synoptic wind flows are transformed by their interaction with the topography and by near-surface processes generating local winds. Numerical weather prediction (NWP) models are increasingly more capable of reproducing such effects, thanks to the constant improvement of their resolution, dynamical cores, and parametrizations. In highly complex terrain like the Alps, fine-scale models are required to resolve the terrain correctly and obtain flows that are close to observations. Despite the improvements in NWP models, a trade-off still exists between the extent of the simulation's domain, its resolution, and the modeled period. In highly complex terrain, given the required resolution, large domains cannot be modeled for long periods. The modeling of many surface processes like snow preferential deposition and transport, however, requires long time series of high-resolution wind fields on large domains (Mott *et al.*, 2010; Vionnet *et al.*, 2021). This is also true for wind-energy potential assessment in such types of terrain, where spatial variability is high.

A solution to the limitation in computational resources is to use NWP models at coarser resolution, succeeded by a statistical downscaling model. The former describe the state of the atmosphere on a coarse topography and the latter uses empirical relationships between these data and another source of information that describes the true nature of the wind. This ground-truth information can have various origins, but it may represent the best partial description of wind flow in real terrain yet. Statistical downscaling is employed in many fields and on many types of data. In complex terrain, wind and precipitation are the prime examples, because of their high spatial variability and importance in surface process modeling. A vast repertoire of techniques is also available. Machine learning is gaining importance in environmental sciences (Hsieh, 2009), and its extension deep learning begins to offer state-of-the-art results (Reichstein *et al.*, 2019).

For downscaling wind in complex terrain, two different approaches are employed, depending on whether the ground truth comes from measurements or from high-resolution NWP models. Philippopoulos and Deligiorgi (2012) trained an artificial neural network (ANN) to predict the hourly wind speed at a certain station given the measurements from other stations in Greece. Dupuy *et al.* (2021) also used an ANN to estimate hourly wind speed at a French station, but from values predicted by the Weather Research and Forecasting (WRF) model with 3-km resolution. Similarly, Goutham *et al.* (2021) used analyses from the European Center for Medium-Range Weather Forecasts (ECMWF, 9 km resolution) as a predictor for random forest or gradient-boosting models to predict wind speed

at 171 French stations. These studies calibrated (trained) point-to-point models that were only capable of predicting wind speed at specific locations. To obtain spatially distributed predictions, Manor and Berkovic (2015) developed a Bayesian-aided selection within a catalog of 0.5-km resolution WRF simulations for northern Israel. The model replicates this high-resolution near-surface wind speed and direction from a coarser 4.5-km WRF dataset. For California, Huang *et al.* (2015) trained a linear model calibrated with WRF at 3 km to downscale the North American Regional Reanalysis (NARR: 32 km, 3-hourly). A deep-learning super-resolution convolutional neural network (CNN) was employed by H ohlein *et al.* (2020) to recreate the ECMWF High-Resolution Forecast (HRES, 9 km) from ECMWF ERA5 (30 km) for southern France. With the exception of the latter study, which integrates a terrain elevation model, these 2D-to-2D methods are constrained to the domain on which they are calibrated, because they do not incorporate descriptors of the topography. Furthermore, their targeted resolutions are not sufficient for highly complex topographies like the Alps. Other types of models like WindNinja (Wagenbrenner *et al.*, 2016; Rios *et al.*, 2018; Hilton and Garg, 2021) and TopoSCALE (Fiddes and Gruber, 2014) are based on physical descriptions of wind-topography interactions and are not bound to the domain of calibration. WindNinja is used frequently for wildfire propagation, but was used recently by Vionnet *et al.* (2021) to model snowdrift. However, given their generic nature, those methods do not benefit from the valuable information of measurements or high-resolution simulations and thus might not perform well in particular terrain under particular conditions.

An interesting line of research took advantage of the vast network of measurement stations in Switzerland to develop point-to-point methods that can generate spatially distributed predictions in the highly complex Swiss Alps. All of them use descriptors of the topography to predict wind speed at the location of the measurement stations. When calibrated, the models can be used at other locations and generate maps of wind speed. Etienne *et al.* (2010) employed general additive models (GAM) for daily maximum wind speed using landform categories and Fischer *et al.* (2015) employed gradient-boosted regression trees and topographic exposure for the same purpose. Foresti *et al.* (2011) and Robert *et al.* (2013), respectively, used support-vector regression on mean wind speed and GAM on monthly values, with various terrain descriptors (slope, difference of Gaussians, and directional derivative). These methods only use topographic descriptors as explanatory variables and are thus limited to temporally static predictions. Without time-dependent inputs like the wind vector from a lower-resolution model, such an approach can only predict one value per location and not a time

series. Using the Consortium for Small-scale Modeling at 2- and 7-km resolutions (COSMO-2, COSMO-7)¹ and the directional-dependent terrain parameter S_x (Winstral *et al.*, 2002), Winstral *et al.* (2017) trained a linear model to generate maps of wind speed at 25-m resolution. As all stations were used for calibration, the model's validation did not reflect the performance at other locations. The results at the stations, however, were promising and showed the importance of considering the distribution of wind speed and not only aggregated metrics. From COSMO-2 and a slope parameter, Helbig *et al.* (2017) proposed a downscaling method that is calibrated from a large catalog of simulations from the Advanced Regional Prediction System (ARPS) run at 30-m resolution. Using stations for validation only, the method slightly outperforms COSMO-2 at exposed sites but is worse at sheltered ones. These two methods offer the resolution and efficiency required for the generation of long time series on large domains mentioned previously. However, they do not consider the change in wind direction when going from coarse to fine resolution and either were not validated spatially or were ineffective on lee sides.

We present a novel 2D-to-point statistical downscaling model: Wind-Topo, first of its kind and based on deep learning, which uses kilometer-resolution NWP model outputs and high-resolution topography to predict near-surface wind speed and direction in highly complex terrain. 261 Swiss measurement stations are used for calibration (training) and 60 different stations are used for validation. In our test case, the 1.1-km resolution COSMO-1 is used as predictor and near-surface wind fields can be downscaled to 50-m resolution. The custom architecture, which relies on multiple CNNs, as well as the custom loss function extract meaningful statistical relationships between the low-resolution atmospheric state surrounding the point of interest, wind measurement, and high-resolution topography. The wind-topography interactions discovered by the model are applied successfully at the validation stations and to a high-resolution grid, and exhibit ridge acceleration, sheltering, and deflection.

2 | METHODS

2.1 | Approach overview

As ground truth, we use wind speed and direction from 321 measurement stations spread across Switzerland from the

¹<https://www.meteoswiss.admin.ch/home/measurement-and-forecasting-systems/warning-and-forecasting-systems/cosmo-forecasting-system.html>

Intercantonal Measuring and Information System (IMIS)² and SwissMetNet³ networks at 7 and 10 m above ground level (m a.g.l.), respectively (Figure 1). From IMIS, 69 “exposed” stations are installed on summits or ridges and 86 “sheltered” stations are in locations that are protected from the main winds (Lehning *et al.*, 1999). The remaining 166 (“other”) stations are uncategorized and are installed in diverse terrain. As we want the model to predict wind speed and direction, and given the noncontinuity of the direction, we use u and v , the horizontal components of the wind vector. We use hourly data from October 1, 2015–October 1, 2018. For the same times, the COSMO-1 analysis dataset describes the local state of the atmosphere surrounding each station, based on the coarse resolution and smoothed topography (slope < 30°) of the model (background of Figure 1). Wind-Topo combines this information with a high-resolution Digital Elevation Model (DEM) to predict the station measurements u and v . In Wind-Topo, “local” means an area of about 21×21 km² (19×19 COSMO-1 pixels) surrounding each station, as depicted in Figure 1. The model's training consists of optimizing the model's parameters such that the predictions \hat{u} and \hat{v} are as close as possible to u and v , using only COSMO-1 and topographic inputs. For training, we use the first and last year of the period mentioned and a selection of 261 stations. The remaining 60 stations and/or the middle year are used to test the model, more precisely its temporal and spatial generalization capabilities.

2.2 | Model inputs

The model's predictors come from either COSMO-1, with the aforementioned spatial extent and resolution, or the high-resolution DEM. For the latter, we combined a 2-m DEM⁴ inside Switzerland and a 1-arcsec DEM⁵ (about 36 m) outside the country. We resampled them such that each COSMO-1 pixel contains exactly 21×21 DEM pixels⁶. As COSMO-1 (on a 0.01×0.0146 lat/lon grid) has a resolution in our whole Swiss domain of 1,113 m in latitude and $1,113 \pm 22$ m in longitude, the high-resolution topographic inputs have a resolution of $53 \times 53 \pm 1$ m. Accordingly, our input topographic patches, covering $21.147 \times 21.147 \pm 0.418$ km, have a resolution of 399×399

²<https://www.slf.ch/en/avalanche-bulletin-and-snow-situation/measured-values/description-of-automated-stations.html>

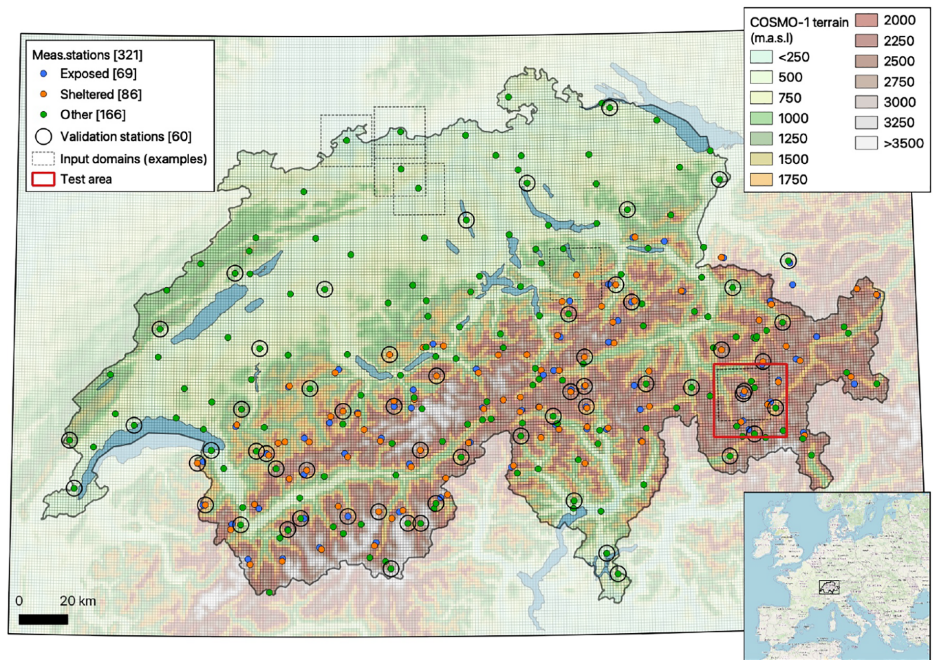
³<https://www.meteoswiss.admin.ch/home/measurement-and-forecasting-systems/land-based-stations/automatisches-messnetz.html>

⁴https://www.swisstopo.admin.ch/en/geodata/height/alti3d.html#technische_details

⁵<https://www2.jpl.nasa.gov/srtm/>

⁶This number of pixels (21×21) is coincidentally the same as the size of the COSMO-1 input domains. There is no link between them.

FIGURE 1 Location of the measurement stations used for training or validation. The background shows the topography used in COSMO-1 and the squares around selected stations represent the spatial extent of Wind-Topo's inputs to predict the wind at those stations. The red and grey squares show the test domain of Section 3.4 and the domain covered in Figure 2, respectively [Colour figure can be viewed at wileyonlinelibrary.com]



pixels. The exact same extent is covered by 19×19 COSMO-1 pixels. The two upper left panels of Figure 2 show an example of such input data from COSMO-1 and DEM.

From COSMO-1, we select 2D fields of horizontal components of the wind vector u_c and v_c (with “c” standing for COSMO) for five vertical layers of the model (terrain-following coordinates) that have average heights of 10, 89, 293, 589, and 1,164 m.a.g.l. We also use the perpendicular deviation of the wind vector from the layer surfaces, denoted w' , which indicates on the five layers where the flow follows the terrain ($w' = 0$), separates upwards ($w' > 0$), or converges towards the ground ($w' < 0$). w' also indicates areas of convection or subsidence. The atmospheric stability is considered via the vertical gradient of potential temperature between the five layers: $\Delta\theta/\Delta h$. The last inputs are the ground surface sensible heat flux q_s and the elevation of the model's terrain z_{COSMO} . Concerning the high-resolution topographic descriptors, elevation z_{topo} and the associated *slope* and *aspect* are employed.

2.3 | New topographic descriptors

Ideally, deep learning could allow an end-to-end approach, where the model learns by itself the operations to apply to the DEM (z_{topo}). However, the rather small variability in topographic inputs (only 321 unique z_{topo} images, one for each local patch) is far from sufficient for obtaining an end-to-end model with good spatial generalization. Such a model could only generate smooth wind fields, lacking small-scale features, and only increased u_c and v_c around

ridges and summits. We thus had to use other topographic inputs that show the wind–topography interactions more directly. Two such descriptors that are well known were initially employed: the topographic position index (TPI), indicating the relative height of a location with respect to its neighborhood, and the maximum upwind slope (S_x) (Winstral *et al.*, 2002), indicating how exposed or sheltered a place is, given a specific wind direction. Both of them share one important parameter: the radius considered for the neighborhood (typically 500–4,000 m). Trained with them, Wind-Topo obtains scores that are almost as good as the ones presented in Section 3. However, the predicted wind fields exhibit a strong dependence on the radius parameter and look like copies of the descriptors, even if several descriptors with various radii are provided. To avoid this problem, we created a set of new topographic descriptors that are free of such parameters and let Wind-Topo decide how to process them. Given its CNN-based nature, Wind-Topo can decide how to integrate them spatially, thus creating an internally controlled and dynamic radius parameter:

$$\begin{cases} E_+ = \max(\sin(\alpha), 0), \\ E_- = \min(\sin(\alpha), 0), \end{cases} \quad \text{where} \quad (1)$$

$$\alpha = \arctan(\tan(\text{slope}) \cos(\delta)), \quad \text{with}$$

$$\delta = \arctan 2(-v_c, -u_c) - \text{aspect}.$$

Figure 2 shows the variables z_{COSMO} , z_{topo} , *slope*, and *aspect* for the input patch centered around the station

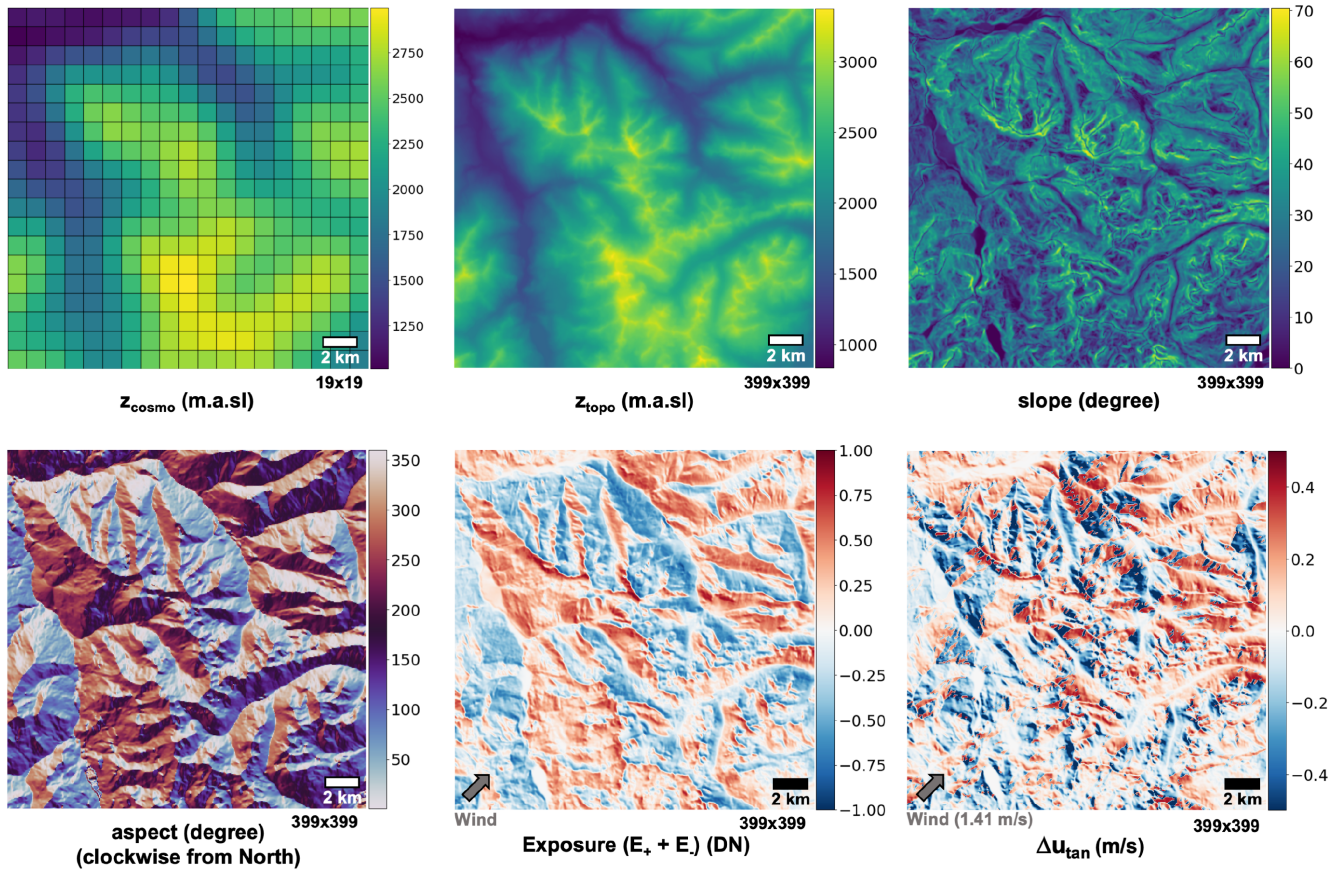


FIGURE 2 Some topographic inputs for a training station (ELA-1). z_{COSMO} is the topography in COSMO-1 and z_{topo} is a 53-m DEM; *slope* and *aspect* are computed from z_{topo} . (E_+ + E_-) reflects places exposed to the wind (red) or sheltered from it (blue). Δu_{tan} shows the offset to apply to u to obtain a flow that is progressively tangential to the terrain as the slope increases, given wind speed and direction [Colour figure can be viewed at wileyonlinelibrary.com]

ELA-1 (Section 3). From *slope*, *aspect*, and the values of u_c , v_c at the center of the patch we generate six topographic descriptors. First, we compute E_+ and E_- (Equation 1) representing respectively how much a location is exposed to or sheltered from the wind, given its direction. The best results were obtained when this direction is computed from the second layer above ground in the COSMO-1 inputs (89 m.a.g.l.). Contrary to S_x , E_+ and E_- do not incorporate a search distance and are purely local. They are computed as the sine of the vertical angle α that the wind vector would have to change to in order to be parallel to the slope while keeping its azimuth. This angle is simply computed from the slope angle and from δ , the angle between the wind direction and the aspect. Splitting this exposure/sheltering into two allows the model to find specific rules for each of them independently. When summed together (as in Figure 2) and integrated over a certain radius, the result resembles S_x . As a CNN cannot multiply some of its inputs together directly, and because our model showed difficulties in learning it, we provide E_+u_c , E_+v_c , E_-u_c , and E_-v_c to the model. The two remaining

descriptors address the challenge of wind deflection: how to change u_c and v_c to obtain a flow that turns when facing a steep slope. We provide the model with the offsets for u_c and v_c , Δu_{tan} and Δv_{tan} , needed to obtain such an effect. Equation 2 shows that the steeper the slope, the more the flow will have to follow the topographic contours (be tangential). This is performed via a rotation of the vector (u_c, v_c) by an angle β . This angle is calculated using the slope angle and δ , which is the same angle as in Equation 1. Using the observations as guidance, Wind-Topo learns how much, where, and when these exposure, sheltering, and theoretical deflections should be employed:

$$\begin{cases} \Delta u_{\text{tan}} = (\cos(\beta) - 1)u_c - \sin(\beta)v_c, \\ \Delta v_{\text{tan}} = \sin(\beta)u_c + (\cos(\beta) - 1)v_c, \end{cases} \quad \text{where} \quad (2)$$

$$\beta = \left(\frac{\pi}{2} - |\delta|\right) \text{sign}(\delta) \sin(\text{slope}), \quad \text{with} \\ \delta \in] -\pi, \pi] \text{ from Equation 1.}$$

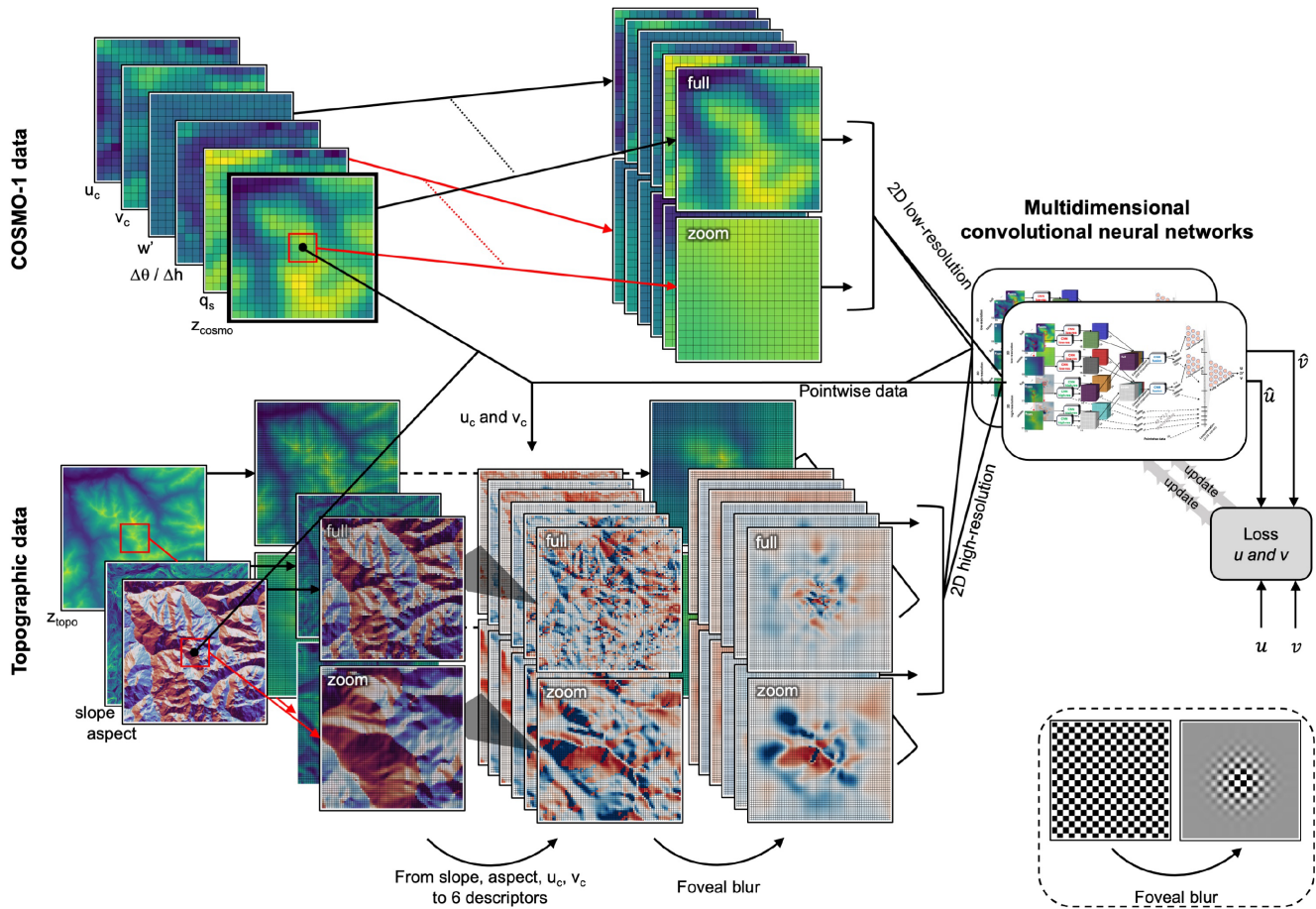


FIGURE 3 Chain of operations to obtain the predicted \hat{u} and \hat{v} at the center of the patch of input data [Colour figure can be viewed at wileyonlinelibrary.com]

2.4 | Data flow

Given the large number of training data (261 stations \times 17,520 timesteps \times n variables), treating the 399×399 inputs is an enormous task (3 TB of data). As Wind-Topo predicts the wind at the center of the input patch, the more relevant information is located close to the center, while the edges of the patch should provide only coarser information. To decrease the amount of information significantly (by a factor of 13) and to help the model focus on high-resolution information towards the center and on progressively coarser information at a distance, we apply two consecutive operations to all topographic inputs (Figure 3). First, each input is split into two: on the one hand the full-extent input is resized to a resolution of 77×77 pixels, and on the other hand a crop of the central 77×77 pixels (a zoom) is extracted. Second, we apply a foveal blur (see bottom right of Figure 3), which smooths the information as we move away from the center. For consistency in the model's architecture, which is described below, the same splitting is applied to the COSMO-1 inputs.

As presented in Figure 3, we have low-resolution COSMO-1 data (full and zoom patches), high-resolution topographic data (full and zoom patches), and values for all those descriptors at the exact location of the station. Time of day and day of the year are also provided as pointwise information via a cosine–sine transformation for continuity. Wind-Topo has two branches: one that predicts u and one that predicts v . Both are independent when making a prediction, but they are trained jointly via a common loss function (Section 2.6) that optimizes their parameters simultaneously. This double architecture gives the best results and allows the model to consider u and v jointly.

A widespread technique in machine learning is to augment the amount and diversity of the training data artificially. For image processing purposes, for example, with CNNs, it is common to rotate, flip, crop, offset, and rescale the images randomly. In our situation, as we only have 261 different input topographies, this augmentation is crucial. 2D data were randomly rotated by one of 72 angles (0–360° every 5°) and all values related to u , v , u_c , v_c , and $aspect$ were transformed in a consistent manner that artificially rotates topography and wind together. To stay

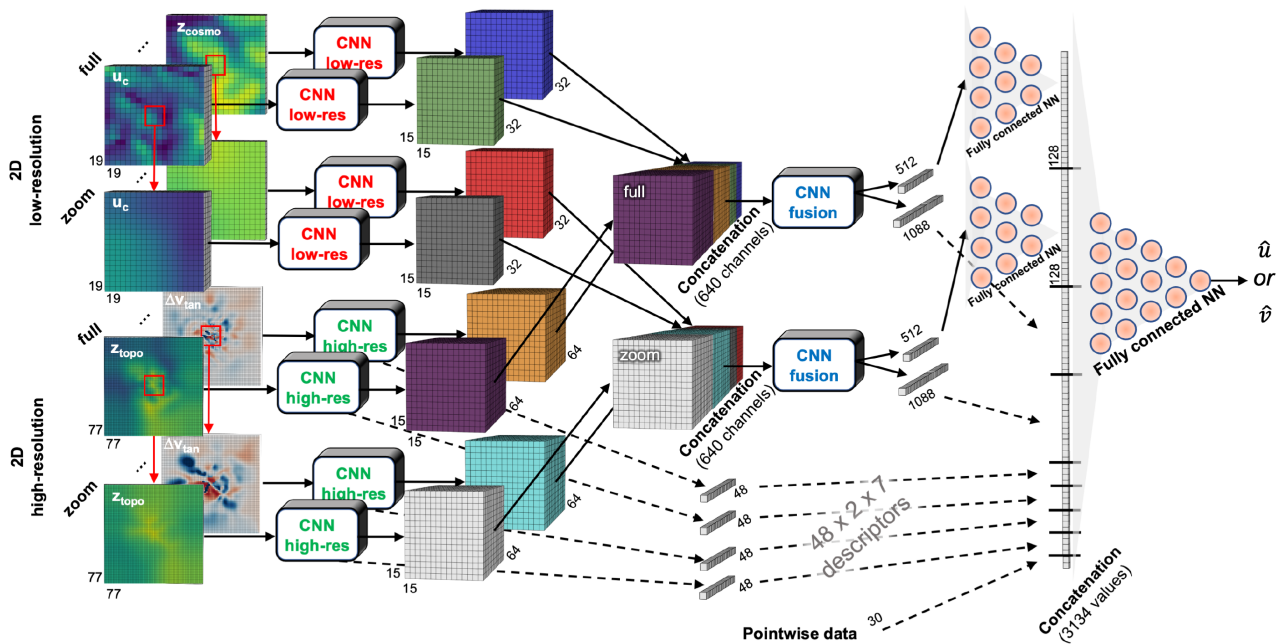


FIGURE 4 Deep-learning architecture of Wind-Topo. Each “CNN low-res” and “CNN high-res” is unique to each input variable. The two distinct “CNN fusion” blocks treat the concatenated tensors with full or zoom extent, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

consistent with wind–topography interactions like preferential deflection to the right due to Coriolis, flip, crop, and rescale were avoided and proved to be ineffective when tried. Offsets of elevation were investigated but did not help. Data augmentation by rotations also allowed us to obtain a model with a good rotational invariance (similar predictions when the inputs are rotated by an angle and the prediction is then rotated oppositely).

2.5 | Deep-learning architecture

Figure 4 shows the constitutive blocks of the model, the data dimensions, and how the various types of data are associated to predict u and v . Each input variable is pre-treated by a dedicated CNN (one CNN for u_c , one for v_c , ...) to generate a tensor. Tensors with similar spatial extent (full or zoom) are concatenated and treated by another type of CNN. As seen in Figures 4 and 5, these latter CNNs and the ones pretreating the high-resolution topographic descriptors generate an additional vector that contains values at the exact center of the patch. This procedure retains important high-resolution information at multiple stages of the CNNs. Classically, the outputs of the CNNs on the concatenated tensors are fed into fully connected neural networks to generate two 128-entry vectors. Finally, all vector data are concatenated and treated by a neural network having a unique output, which is \hat{u} or \hat{v} depending on the model’s branch.

Figure 5 details the building blocks of the deep-learning architecture. The CNNs have a classical alternation of convolution layers and pooling layers in order to extract increasingly larger and higher-level features from each piece of input data and from the concatenated data. Special attention was given to the size of the inputs and kernels, as well as to the strides, to avoid padding on the edges and lateral shifts of information. As Wind-Topo gives a prediction for the exact patch’s center, it is essential that each operation leads to an odd number of rows and columns and that the kernels fall centered on the central pixel. The sigmoid linear unit (SiLU) function ($\text{silu}(x) = x\sigma(x)$, where $\sigma(x) = 1/(1 + e^{-x})$ is the logistic sigmoid: Hendrycks and Gimpel, 2016) was the best activation function, outperforming the standard rectified linear unit (ReLU) function slightly, most likely because of its continuity (important for a regression problem). A particularity of the CNNs treating the high-resolution topographic data is the addition of a directional mask after the first activation. This mask, which sets to zero the area that is downwind from the patch’s center, has two advantages: (i) it helps the model understand the notion of wind direction and how to deal with windward and leeward information, and (ii) it augments the number of distinct topographic inputs. As mentioned previously, central values (red) are extracted before each pooling layer and before several convolution layers in order to retain low-level high-resolution information about the center of the patch.

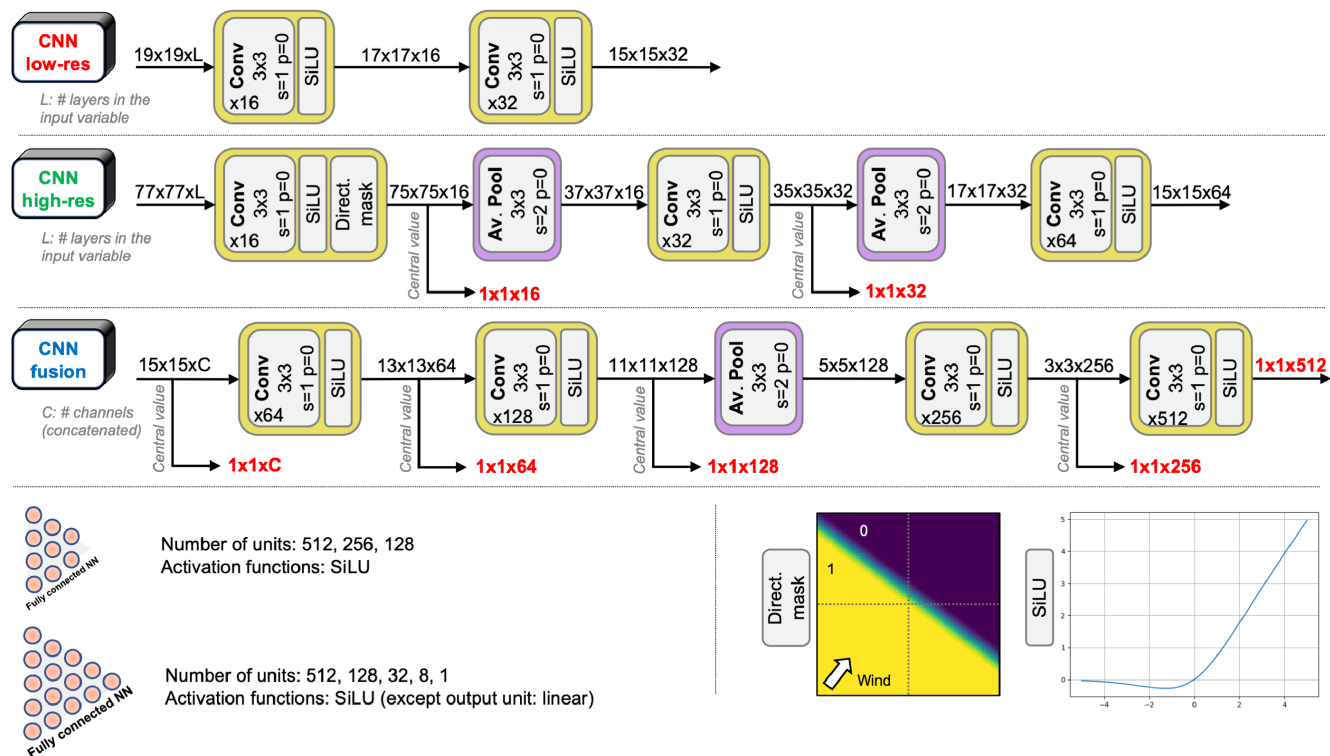


FIGURE 5 Constitutive blocks of deep learning in Wind-Topo. The convolution blocks (yellow) indicate kernel size (3×3), stride s , padding size p , and number of filters (e.g., $\times 16$). The average-pooling blocks (purple) depict the same information. Vectors of central information (red) are extracted at various places. The first convolution layer of “CNN high-res” has an additional directional mask based on the wind direction from COSMO-1 (e.g., mask in lower right corner). The two types of fully connected neural network are detailed in the lower left corner [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Hyperparameters used to train the model

Parameter	Value
Optimizer	ADAM
β_1, β_2	(0.9, 0.999)
Initialization	Xavier, uniform, gain=1
Learning rate	5×10^{-6} , no weight decay
Early stopping	10th epoch
Batch size	128

ADAM: (Kingma and Ba, 2015).

Xavier: (Glorot and Bengio, 2010)

Training the model consists of optimizing the values of the model’s parameters (weights and biases for neurons, filters in a CNN). One needs to adopt a strategy to do so, in order to reach the best performance on the training and validation sets. The various elements that compose this strategy are generally referred to as the hyperparameters of the model. For Wind-Topo, an iterative procedure led to the hyperparameters presented in Table 1.

The version of Wind-Topo presented here is the result of extensive experimentation that aimed at finding the

best architecture, input data, and training procedure in a large design space. In Appendix S1, we present several model variants and their performance. This ablation study shows the relative importance of various design decisions by removing one model component at a time. It appears that the current design is not yet optimized with respect to model complexity and that some building blocks, despite improving performance slightly, significantly increase computational requirements. Future research will investigate the trade-off between model complexity and downscaling performance in more detail.

2.6 | Loss function

The design of an appropriate loss function was critical to obtain the current performance of Wind-Topo. Using any standard loss function like mean-square or absolute error (MSE, MAE) led to distributions of \hat{u} , \hat{v} and the associated horizontal wind speed \widehat{vel} that were too narrow. The model chose to decrease the range of its predictions, favoring values near the mean value. This is a well-known problem in

statistical downscaling (Winstral *et al.*, 2017).

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \tau_i [(\hat{u}_i - \beta_i u_i)^2 + (\hat{v}_i - \beta_i v_i)^2],$$

$$\beta_i = \frac{\epsilon + vel_i}{\epsilon + \widehat{vel}_i}, \quad \tau_i = \begin{cases} \tau & \text{if } \widehat{vel}_i \geq vel_i, \\ 1 - \tau & \text{if } \widehat{vel}_i < vel_i. \end{cases} \quad (3)$$

The countermeasure was to compose the loss function \mathcal{L} in Equation 3, which is essentially a MSE over a batch of training data of size N , with a scaling by β_i of the target values u_i and v_i based on the overestimation or underestimation of the predicted velocities \widehat{vel}_i . Additionally, the term τ_i , inspired by the pinball loss function (Takeuchi *et al.*, 2006), allowed the model to remove the consistent negative mean bias error (MBE) of \widehat{vel} . Parameters ϵ , controlling the strength of the scaling, and τ were tuned iteratively to their respective optimal values of 4 and 0.425. A lower value of ϵ led to an unstable behavior and incorrect distributions, while larger values led to squeezed distributions. The current value of τ allowed us to obtain an unbiased model on the test set. The quality of the distributions of predicted wind speed is essential for many applications of high-resolution wind fields. This loss function is the result of an intense effort to find a differentiable function that guides the gradient-descent algorithm of the model's optimization efficiently towards distributions that closely match the measured ones, while obtaining very low MAE and MBE and high Pearson correlation coefficients between predictions and measurements. These three scoring metrics reflect important qualities of the predictions. A low MBE guarantees no bias between predictions and observations, while a simultaneous low MAE and high correlation indicate that the model is accurate for a large range of observations. The development of Wind-Topo was guided by using these three scoring metrics on the validation set, and Section 3 focuses on them.

2.7 | Making predictions

Wind-Topo was trained to make a point prediction from centered patches of COSMO-1 data and a high-resolution DEM covering exactly 21.147 km. Thus, when Wind-Topo is used to generate a wind field on, for example, a 100-m grid, this is performed in a point-by-point manner. Each predicted grid point is an independent output of the model, to which we provide the centered patches of data. In this 100-m case, COSMO-1 needs to be resampled every 100 m to avoid 11 consecutive grid points receiving the same COSMO-1 inputs, which would lead to a chessboard-like pattern in the resulting high-resolution wind field. This

subpixel resampling is performed efficiently using convolutions with Lanczos kernels (Burger and Burge, 2009). If the same method is applied to the topographic data as well, it is possible to generate wind fields at a higher resolution than the 53-m DEM input data. However, as the model was trained on this particular DEM, higher resolutions lead to smooth outputs with no additional information. In our case, wind fields generated at a 50- to 100-m resolution contain the highest level of detail.

The complex pipeline needed to obtain all the required input data for every grid point can be performed by the CPU, in parallel with the predictions performed by the GPU. In our case, for Wind-Topo coded with PYTORCH (Paszke *et al.*, 2017),⁷ it is possible for any domain size to obtain a perfect parallelization of these two tasks, without noticeable overhead and with a GPU running constantly at full capacity. As the model is pointwise and time-independent, it is easily parallelizable on multiple GPUs and/or machines with, for example, one GPU computing one part of the domain or a subperiod and another GPU computing the rest. On a NVIDIA RTX2080Ti GPU, one prediction takes 0.6 ms and a 300 × 300 domain requires 54 s. This is 30% faster than on a previous-generation GTX1080Ti, which indicates that Wind-Topo would be even faster on the latest-generation high-performance GPUs.

3 | RESULTS

This section shows the overall performance of Wind-Topo in highly complex terrain and quantifies its aptitude for capturing subgrid-scale wind-topography interactions like ridge acceleration, sheltering, and deflection. After training Wind-Topo on 261 stations and years 1 and 3, we tested it on separate datasets to assess its performance at new locations and/or over a new time period: test set 1 (60 stations, years 1 and 3), test set 2 (261 stations, year 2), and test set 3 (60 stations, year 2) being the most informative because it reflects the spatio-temporal generalization. In this section, we will evaluate Wind-Topo first quantitatively and then qualitatively by discussing some predicted wind fields. From now, when COSMO-1 is mentioned, it should be understood that a bicubic interpolation was applied at the desired locations. Importantly, the 60 test stations were selected by an algorithm that ensures an equitable evaluation. First, all stations were tagged with an eight-dimensional vector consisting of (1) mean measured wind speed, (2) MAE (of COSMO-1 with respect to measurement), (3) MBE, (4) correlation, (5) normalized MAE

⁷<https://github.com/pytorch/pytorch/tree/v1.8.1>

TABLE 2 Performance of COSMO-1 (C) and Wind-Topo (WT)

Station (\overline{vel})	Model	Corr. (DN)	MBE ($\text{m}\cdot\text{s}^{-1}$)	MAE ($\text{m}\cdot\text{s}^{-1}$)	nMAE (DN)	MAEdir (deg)
Exposed ($3.43 \text{ m}\cdot\text{s}^{-1}$)	C	0.51	-0.25	1.83	0.56	41.3
	WT	0.66	-0.05	1.49	0.45	32.5
Sheltered ($1.67 \text{ m}\cdot\text{s}^{-1}$)	C	0.42	1.90	2.27	1.43	67.0
	WT	0.45	-0.04	1.03	0.63	58.1
Other ($2.46 \text{ m}\cdot\text{s}^{-1}$)	C	0.59	0.18	1.34	0.61	50.5
	WT	0.64	-0.16	1.16	0.51	47.5
Alps ($2.35 \text{ m}\cdot\text{s}^{-1}$)	C	0.50	0.72	1.77	0.90	56.4
	WT	0.55	-0.07	1.21	0.56	50.1

Note: Each value is the average of the scores at each test station, computed for the test period. The left column gives the average measured wind speed for each station category.

Abbreviations: Corr., Pearson correlation coefficient; MAE, mean absolute error; MBE, mean bias error; nMAE, normalized MAE (MAE at a station divided by its average measured wind speed); MAEdir, mean absolute error of wind direction.

(MAE at a station divided by its average measured wind speed), (6) elevation, (7) latitude, and (8) longitude. Then, in an iterative manner, the algorithm splits the stations randomly into training and test sets until there is a good representation of this eight-dimensional space in both sets and a similar performance of COSMO-1. Concerning the separation of training and test times, our investigation showed that a random split cannot reflect the temporal generalization: it is easy for a model trained on hours h and $h + 2$ to predict hour $h + 1$. Consequently, we kept the complete intermediate year for validation. In this section, we employ the same colors as in Figure 1 for stations that are exposed (blue), sheltered (orange), and uncategorized (“other”, green).

3.1 | Aggregated scores

Table 2 shows various scoring metrics for COSMO-1 and Wind-Topo on test set 3. Each score is the average of the scores at the stations belonging to that category. This is more meaningful than computing the scores on all the data, as, for example, correlations computed on all the data are significantly higher but are potentially misleading.

We can observe the importance of the station’s classification in the left column: the average wind speed at exposed sites is more than twice that at sheltered sites. Given its resolution, COSMO-1 cannot discriminate between them and has a large positive bias ($1.9 \text{ m}\cdot\text{s}^{-1}$) at sheltered sites. Surprisingly, the highest wind speeds are predicted there (Krutz *et al.*, 2018). Nevertheless, its performance is quite high at exposed locations and in flat lands and large valleys (a considerable part of the “other” stations). Wind-Topo improves COSMO-1 significantly for all

metrics and at all station types. The correlation is improved the most at exposed and other sites, while the large bias at sheltered sites is completely removed, leading to a much lower MAE. The normalized MAE (nMAE) allows us to compare all types of station and again shows the lack of performance of COSMO-1 at sheltered sites and the capability of Wind-Topo to correct it. Finally, Wind-Topo also corrects wind directions, as expressed by the MAE of wind direction (MAEdir), computed from events with a wind speed greater than $1 \text{ m}\cdot\text{s}^{-1}$. As Wind-Topo was designed for complex terrain, the lower line in Table 2 provides overall scores for the 44 test stations located in the Alps.

The scores for Wind-Topo were obtained from predictions of the trained model. Figure 6 depicts some of them during the training. After each epoch (pass through the training set), we evaluated the model on the four datasets. Correlation, MAE, and MAEdir improved progressively until epoch 9, the model state of which was used for all predictions in this work. Later on, a typical overfitting is observed, where the spatial generalization decreases. Interestingly, there is no overfitting with respect to the temporal generalization. When training the model further (not shown), the scores on test set 2 keep improving (MAE $< 1 \text{ m}\cdot\text{s}^{-1}$, correlation > 0.8). Temporal generalization was easily obtained: many (simpler) models can give accurate predictions on test set 2. In other words, the corrections to COSMO-1 can be learned easily for a certain period and applied successfully to new periods. Furthermore, a small number of timesteps is required. All results presented are based on Wind-Topo trained with only half the training times (random selection), and no improvement was observed using all times. The strong temporal generalization offers a different application: if a short measurement

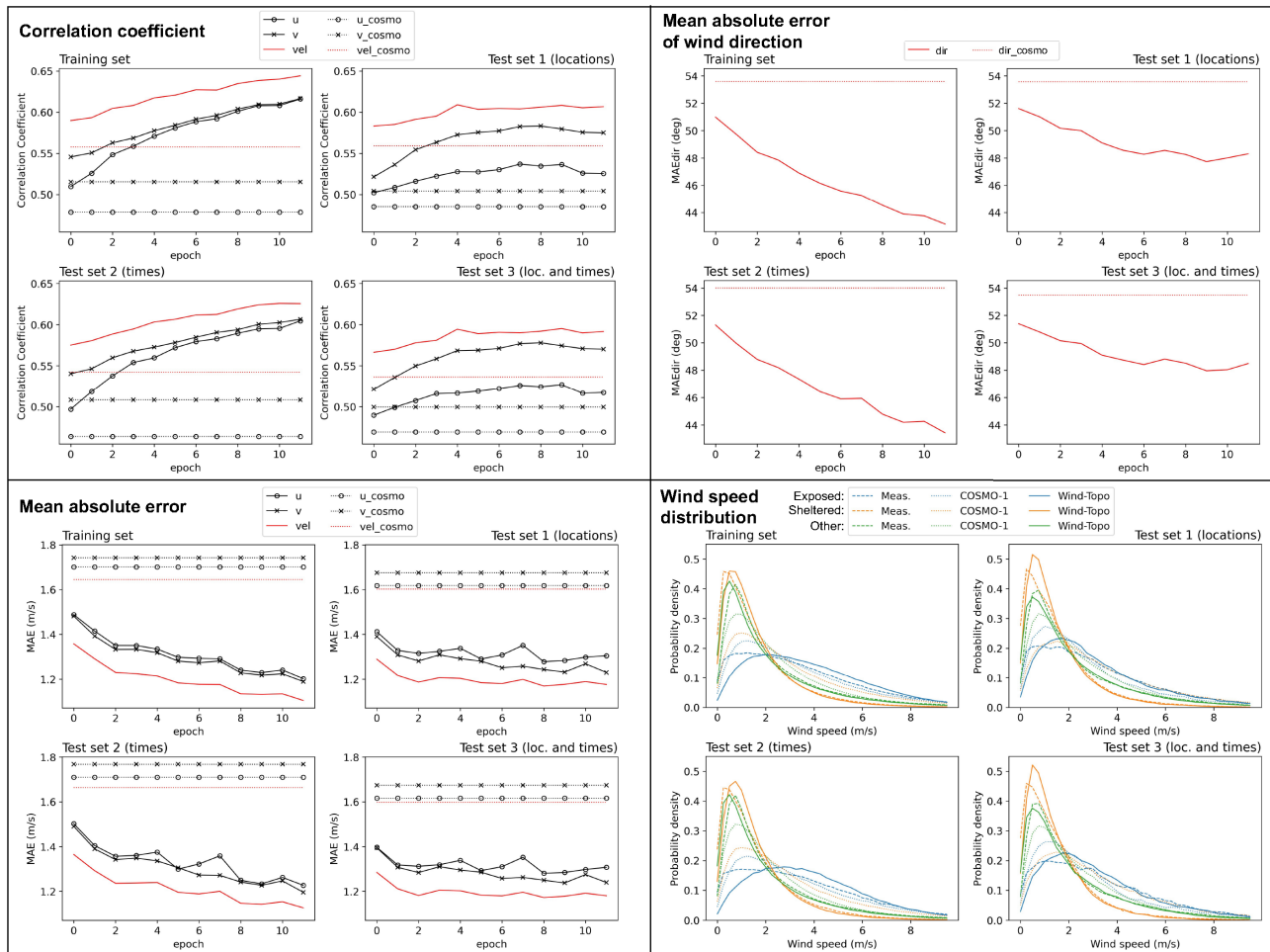


FIGURE 6 Quality of the predictions of Wind-Topo compared with COSMO-1 throughout the training phase (one epoch is one pass through the entire training set). For each quality assessment, the upper left panel corresponds to the behavior of Wind-Topo on the training set, while the right panels correspond to the test stations and the lower panels to the test period. Scores for wind speed (vel) and direction (dir) are computed from u and v and are depicted in red. The lower right figure shows, with various colors, the distributions of wind speed for the three station categories. Various lines styles are used to depict those distributions from measurement, COSMO-1, and Wind-Topo [Colour figure can be viewed at wileyonlinelibrary.com]

campaign of surface wind is performed somewhere, the collected data can be used to train a model like Wind-Topo, which can generate longer “synthetic” time series covering the same period as the NWP model. The lower right panel of Figure 6 shows the distributions of measured and predicted wind speeds for the three station types. We observe again the lack of discrimination of COSMO-1 between exposed and sheltered sites. It is, however, clear in the measurement, and Wind-Topo can reproduce it well. The distributions of predicted wind speed match the observed ones accurately for all station types and for the four datasets.

The “early stopping” procedure presented above, in association with the chosen learning rate, provides the best scores and distributions. Other regularization techniques like weight decay did not show any supplementary advantage. On a RTX2080Ti GPU, the 12 epochs of Figure 6 required 43 hr.

3.2 | Disaggregated scores

We can assess Wind-Topo further by looking at the scores at each station. Figure 7 shows such a disaggregated view for the MAE. Appendix S1 provides similar plots for correlation, MBE, nMAE, and MAEdir. Each segment represents a station, with its extremities being the scores on training and test periods. The test stations are circled in black and are the ones we should focus on. Any point located below the black curve indicates that Wind-Topo has a lower MAE than COSMO-1. This is the case for all exposed and sheltered stations. Wind-Topo is, however, slightly worse at some “other” stations and significantly worse at one. The clustering of the station categories is remarkable. Wind-Topo strongly reduces the MAE at the sheltered sites with high COSMO-1 MAE. This is also the case at exposed sites, albeit less pronounced. The reduced performance at the “other” stations is due to the already

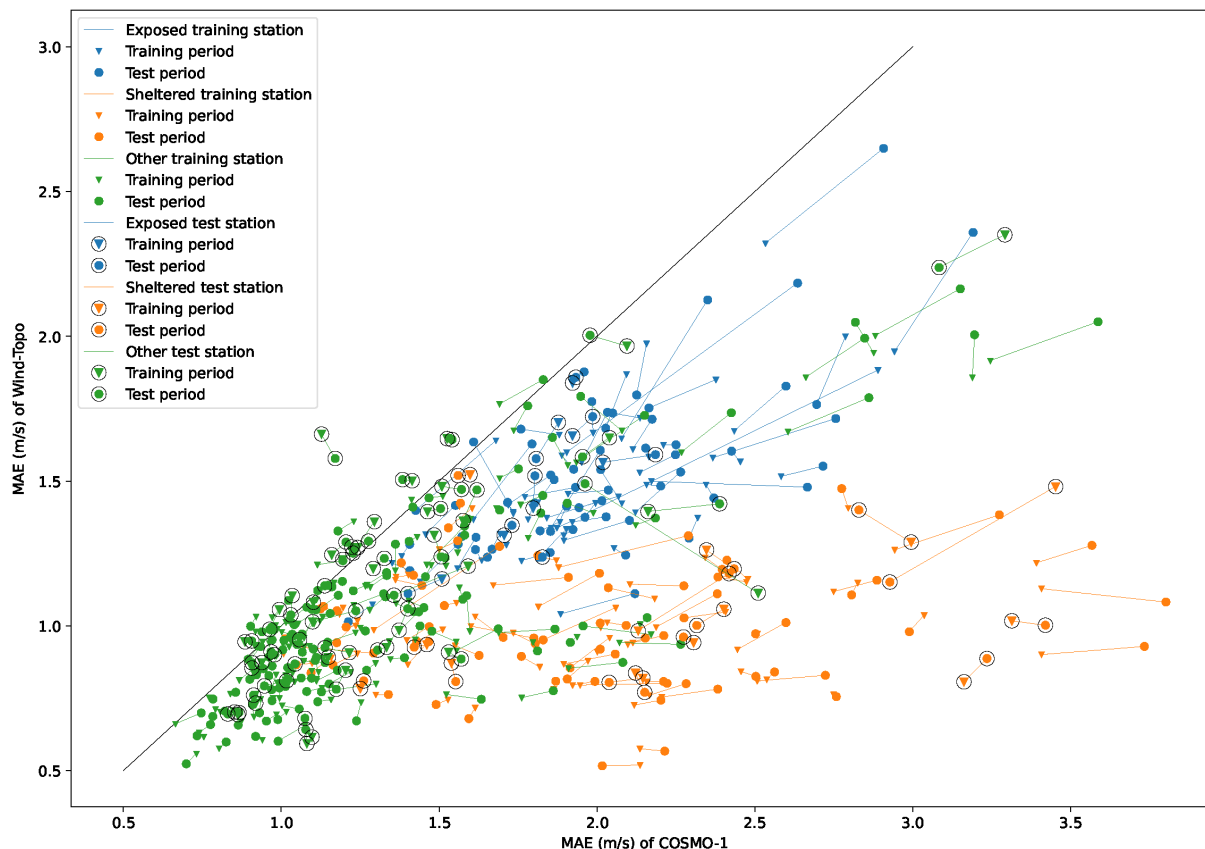


FIGURE 7 Mean absolute error of Wind-Topo at all 321 stations compared with COSMO-1. Each segment corresponds to one station. The color indicates the category of the station. The extremities of each segment show the MAE for the training period (triangles) and the test period (circles). The 60 test stations are circled in black. Any point below the black line shows an improvement from Wind-Topo [Colour figure can be viewed at wileyonlinelibrary.com]

good performance of COSMO-1. The orientation of the segments indicates how the MAE in COSMO-1 changes from one period to another and how Wind-Topo reacts to it. An orientation parallel to $y = x$ reflects that Wind-Topo follows the variation of score in COSMO-1. An upward deviation from this orientation (triangle to circle) shows that Wind-Topo suffers from a lack of temporal generalization and becomes (relatively) worse for the test period. This is almost never the case and the opposite behavior dominates.

The “outlier” mentioned above is an interesting test station located on the smooth ridge of a hill (700 m above the plains). Its topography is well represented in COSMO-1, which performs well with almost no MBE ($0.2 \text{ m}\cdot\text{s}^{-1}$) and a low MAE ($1.2 \text{ m}\cdot\text{s}^{-1}$). Wind-Topo overestimates the wind speed there (MBE of $1.2 \text{ m}\cdot\text{s}^{-1}$), which leads to a higher MAE ($1.5 \text{ m}\cdot\text{s}^{-1}$). Surprisingly, at another (training) station located 11 km from there and 450 m higher on the ridge, COSMO-1 behaves very differently. It underestimates the wind speed ($4.6 \text{ m}\cdot\text{s}^{-1}$) compared with measurement ($7.7 \text{ m}\cdot\text{s}^{-1}$). This large bias is almost entirely corrected by Wind-Topo (MBE of $-0.9 \text{ m}\cdot\text{s}^{-1}$). Being in the training set, this performance does not necessarily

reflect appropriate corrections. However, it seems that the learned corrections of COSMO-1’s underestimation in this type of topography are applied to our nearby “outlier”. Such topographies are not common in our dataset and Wind-Topo would certainly benefit from more training stations, located in more diverse terrain.

Figure 8 provides a temporal disaggregation of the scores for each month of the test period. The 0.25, 0.5, and 0.75 quantiles, computed for the 60 test stations, are represented in bright colors for Wind-Topo and faint colors for COSMO-1. The upper panel shows that the winter period is the windiest and that COSMO-1 predicts similar wind speeds at exposed and sheltered sites. This large positive bias (MBE panel) is always corrected by Wind-Topo. The negative bias occurring at exposed sites in winter is also corrected, as well as the slight positive bias at the “other” stations in summer. The panel on correlation shows the significant improvement from Wind-Topo at exposed sites, especially in early summer, when thermally driven flows become active. At the “other” stations, Wind-Topo follows closely, but slightly surpasses, COSMO-1. The same happens at sheltered sites, except in July–September, when the correlation is

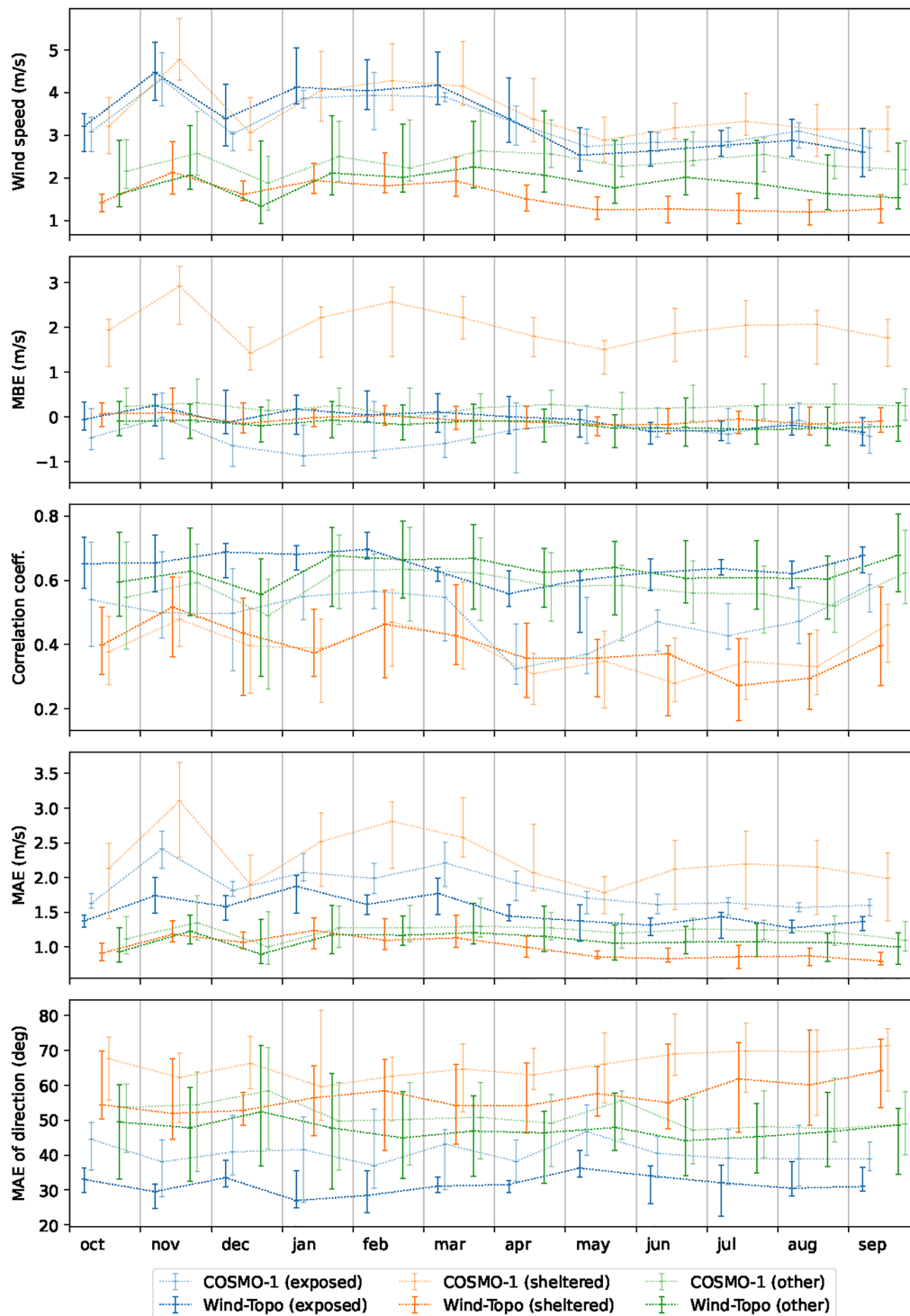


FIGURE 8 Each panel shows monthly average values for the three types of test station (colors). Those values are the 0.25 quantiles (lower extremities of the vertical segments), 0.5 quantiles (marks within the segments), and 0.75 quantiles (upper extremities) of the monthly values for each station. Values for Wind-Topo are depicted with bright colors, while values for COSMO-1 are fainter [Colour figure can be viewed at wileyonlinelibrary.com]

slightly decreased in Wind-Topo, despite a low MAE. We should note the low average wind speed for this period ($1.2 \text{ m}\cdot\text{s}^{-1}$) and the corresponding large MAE_{dir} from COSMO-1 (68°). Wind-Topo can lower it to 60° , but at a cost of a lower correlation. Finally, the MAE of

wind speed and direction are systematically lower with Wind-Topo. This is true for the 0.5 quantile and almost always true for the other quantiles, showing that the downscaling performs well at almost all stations and for all seasons.

3.3 | Selected stations

For a last quantitative analysis, we selected three test stations in the test area of Figure 1: ELA-1, ELA-2, and SAM-0. Indices 1, 2, 0 stand for exposed, sheltered, and other, respectively. We can first look at the wind-rose diagrams for those stations (Figure 9), generated with data from measurements, COSMO-1, and Wind-Topo. For information, their locations are depicted in Figure 10. We can observe that the wind direction in Wind-Topo is strongly influenced by COSMO-1. This is especially the case at ELA-1, where Wind-Topo can only correct the westerly winds from COSMO-1 partially. At SAM-0, the

situation is reversed, with the COSMO-1 direction being accurate while Wind-Topo rotates it slightly. This can be explained by the peculiar location of this station (intersection of several valleys) and it reflects a weakness of Wind-Topo. At ELA-2, Wind-Topo successfully corrects wind speed and direction. These three stations illustrate the range of potential corrections depending on station characteristics. At exposed sites like ELA-1, COSMO-1 performs reasonably well but Wind-Topo can still improve the flow. At sheltered locations like ELA-2, the coarse resolution of COSMO-1 cannot reflect the sheltering and this leads to a strong overestimation of wind speed and wrong wind direction. Wind-Topo shows a clear ability to lower

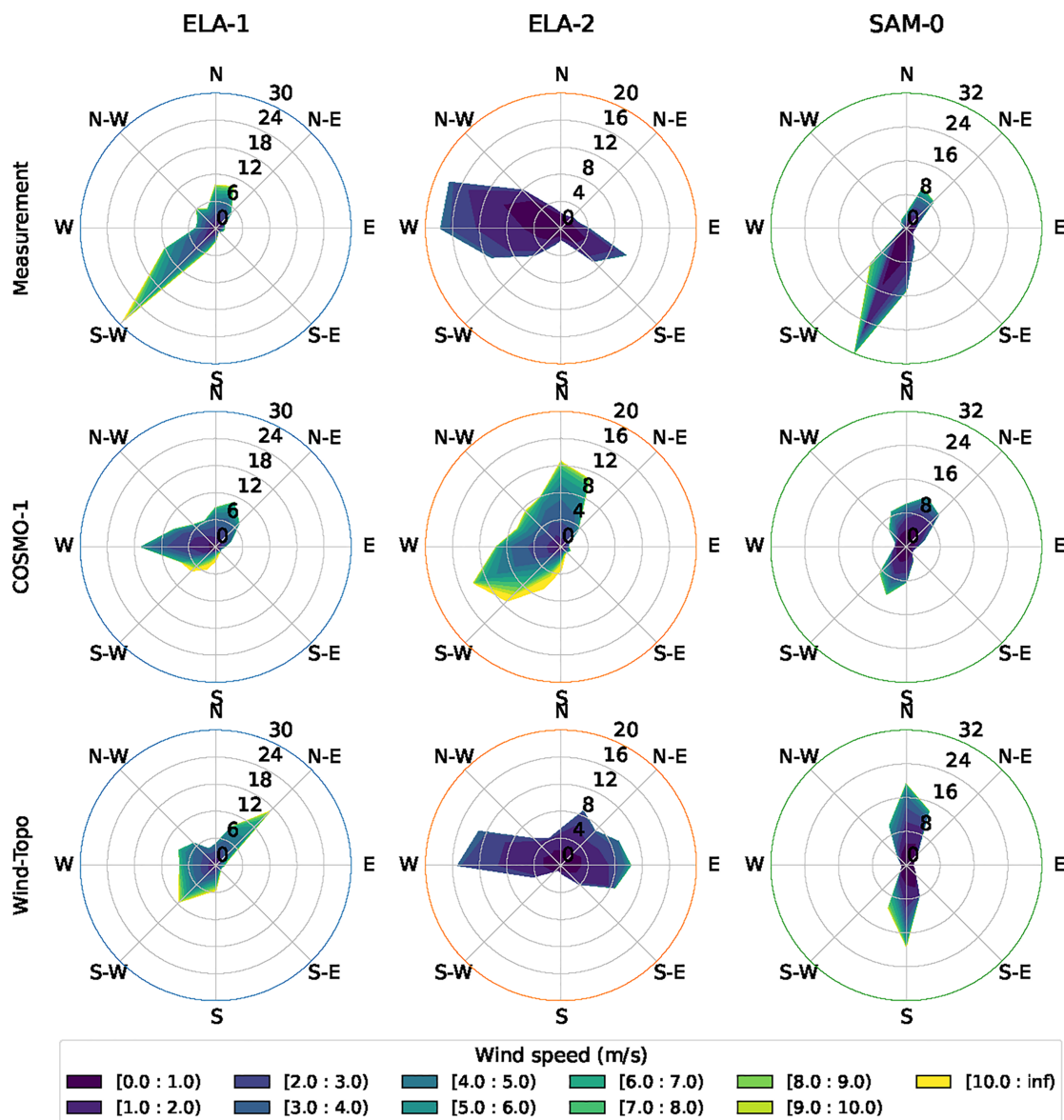


FIGURE 9 Wind-rose diagrams for the three selected test stations during the test period from measurement, COSMO-1, and Wind-Topo. The color indicates wind speed, while the distance from the center indicates the probability in percent that the wind comes from one of 16 directions. To ease the comparison, the radial axis has the same extent for the three data sources [Colour figure can be viewed at wileyonlinelibrary.com]

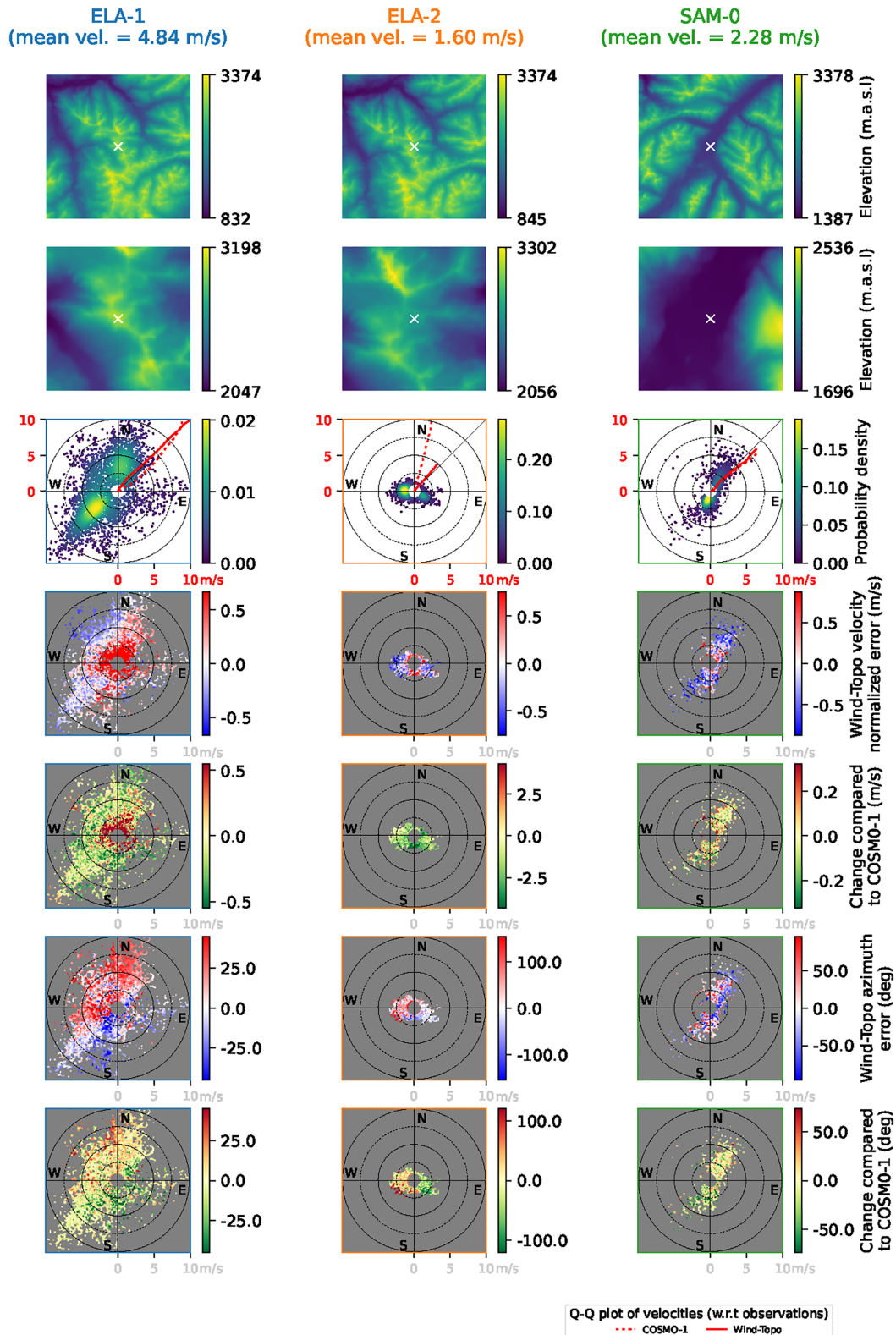


FIGURE 10 Wind polar plots for three selected test stations. The upper two rows show the topography used by Wind-Topo (top: full domain, bottom: central crop). The lower rows are all constructed in the same way: polar plots of a desired quantity by wind speed and direction. For example, strong northwesterly winds are located in the upper left corner. Events with wind speed smaller than $1 \text{ m}\cdot\text{s}^{-1}$ are not depicted. The colors indicate the average value for all events with a given wind speed and direction. Row 3 depicts the probability density, on which quantile–quantile plots for Wind-Topo and COSMO-1 (of wind speed with respect to measurements) are superimposed. Rows 4 and 5, respectively, depict the relative error in wind speed from Wind-Topo (blue: underestimation, red: overestimation) and how this error changed between COSMO-1 and Wind-Topo (green being an improvement). The lower two rows are equivalent for wind direction (azimuth) [Colour figure can be viewed at wileyonlinelibrary.com]

the wind speed to match observations better and to correct the wind direction. Finally, at locations where the topography is already well described at COSMO-1 resolution, Wind-Topo is not able to add relevant information and might even decrease the quality of the prediction slightly.

In Figure 10 we present a new type of plot that reveals how a model like COSMO-1 or Wind-Topo performs (with respect to observations) across diverse wind conditions. The first two rows depict the local topography of the three selected stations (full and zoom patches of input data), while the lower rows, coined “wind polar plots”, are polar plots by wind speed and direction and can be seen as a variation of wind-rose diagrams. The color of each point represents a certain quantity like the bias in wind speed, which is the average value for all events with such wind speed and direction. For example, strong northwesterly wind events are located in the upper left corner. Additionally, we superimposed on the third row, which shows the probability density, red curves corresponding to quantile–quantile (Q–Q) plots for Wind-Topo and COSMO-1 (of wind speed, with respect to measurement). For ELA-1, the Q–Q plot shows a small underestimation of COSMO-1 for all but high wind speed. Wind-Topo provides an almost perfect correction, except for a small overestimation at low wind speed.

The fourth row confirms the small bias at low wind speed, but no clear dependence on wind direction. The row below depicts the improvement (green) or worsening (red) of Wind-Topo compared with COSMO-1. Except for wind speeds below $2 \text{ m}\cdot\text{s}^{-1}$, the improvement is consistent across all wind speeds and direction. The bottom two rows offer a similar analysis for the error in wind direction. Wind-Topo has a small overestimation of the azimuth when the wind is southwesterly to northeasterly, and an underestimation otherwise. Compared with COSMO-1, the improvement is mixed and presents the same split as just mentioned. For brevity, the reader is encouraged to observe the wind polar plots for other stations with various topographies. Appendix S1 provides the plots for all 60 validation stations. Looking simultaneously at the probability density plots and the plots for other metrics, as well as topography, helps us understand how COSMO-1 performs in the most frequent wind situations and how Wind-Topo reacts. Wind-Topo seems to perform well across a variety of terrain and wind speed and direction, but it cannot, however, introduce small-scale thermally driven flows that are not present in COSMO-1. At some sheltered sites, located on slopes exposed to the sun, Wind-Topo fails to recreate the observed diurnal slope flows when they are not present in COSMO-1. A decomposition of the wind polar plots for day and night events in summer (not provided) reveals it even more clearly. Ongoing effort is addressing this weakness and we expect Wind-Topo to be able to introduce at

least some of those locally generated flows in the near future.

To conclude our analysis for the three stations selected, Figure 11 details the predictions of the two models throughout the test year and for each hour of the day. The histograms of wind speed for the whole year reflect the typical behavior of COSMO-1 at two nearby exposed and sheltered stations (ELA-1, ELA-2): almost identical, which does not correspond to measurements. As seen previously, this is corrected by Wind-Topo. The daily patterns of wind speed are quite different from site to site and for different times of the year. SAM-0 has typical, thermally driven flows in the mid-afternoon, which are stronger in summer. COSMO-1 captures them very well and Wind-Topo replicates them. The measurements at ELA-1 show peculiar wind patterns: wind is steadier and stronger in winter, and is significantly reduced in the middle of the day in summer due to complex interactions between the boundary layer and the free atmosphere. COSMO-1 predicts the opposite, with an increase of wind speed in the middle of the day in summer. It is remarkable that Wind-Topo can reproduce the observed patterns and does not simply copy COSMO-1 at this exposed station, which is located on a type of terrain where COSMO-1 normally performs quite well.

This quantitative analysis displayed many strengths of Wind-Topo, which can (a) distinguish accurately between exposed and sheltered locations and reduce the biases resulting from a simple interpolation of COSMO-1 to a higher resolution, (b) generate better wind-speed distributions in complex terrain, and (c) capture complex small-scale wind–topography interactions like ridge acceleration, sheltering, and deflection. The model is not perfect, as it fails to introduce subgrid-scale thermally driven flows and might introduce errors in locations where COSMO-1 performs well. Given the current level of performance, based on only 261 training stations, we are confident that more diverse training data (ground truth at other locations) would benefit Wind-Topo strongly.

3.4 | Generated wind fields

We can finally look at some examples of generated wind fields for the test area in Figure 1. COSMO-1 was down-scaled for all hours of the test year to a 100-m grid covering this $30 \times 30 \text{ km}^2$ domain, with a bicubic interpolation and with Wind-Topo. The latter required 5.5 days on a RTX2080Ti GPU. Figure 12a shows an example of an interpolated COSMO-1 wind field and the underlying (interpolated) model terrain. Figure 12b is the 100-m DEM, and Figure 12c shows the prediction from Wind-Topo with wind speed enhanced on ridges and exposed slopes and reduced on sheltered slopes. It is not a pale copy of

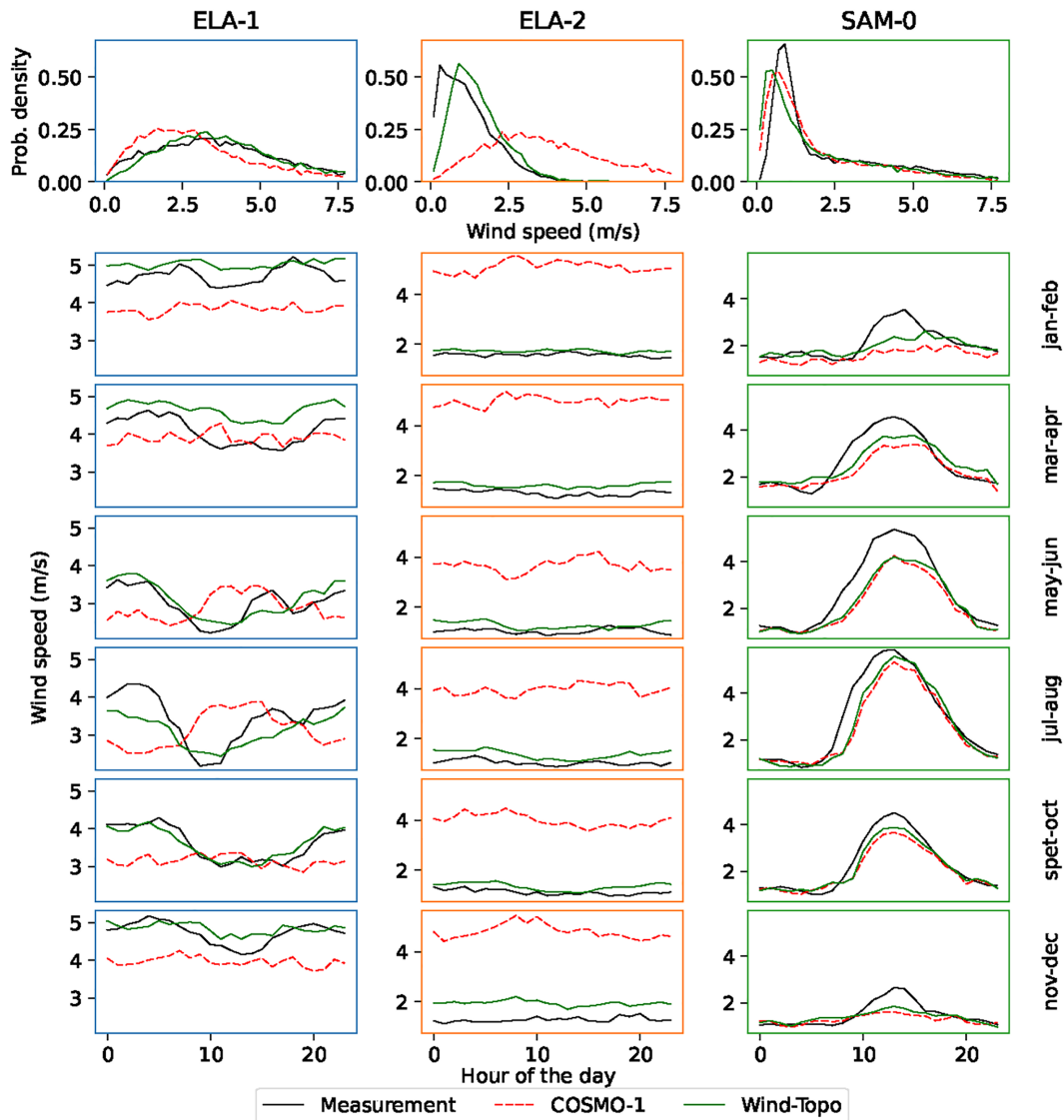


FIGURE 11 Upper panels: distributions of wind speed for the test period from measurement (black), COSMO-1 (red), and Wind-Topo (green) for the three test stations selected. With the same colors, the lower panels show, for each pair of months (in the test period), the average wind speed (y-axis) for each hour of the day (x-axis) [Colour figure can be viewed at wileyonlinelibrary.com]

COSMO-1 with a simple scaling by elevation or exposition. A careful observation (full-size panels available in Appendix S1) indicates the presence of flow deflection and recirculation areas on some lee sides. COSMO-1 shows such effects on a much larger scale. Wind-Topo can even remove them in some areas, before introducing them in others. We should also note the continuity in wind speed in the generated field, despite the point-by-point method. This is caused by the large overlap of input data between neighboring points. To visualize the flow deflection and continuity of wind direction better, Figure 12d depicts the wind azimuth with a cyclic color scale and associated streamlines.

Figure 12e,f provides an insight into the entire down-scaled dataset. Figure 12e shows the average wind speed

and Figure 12f depicts the distributions at every grid point. A Weibull distribution is fitted to the histogram of wind speed at each grid point, and the corresponding parameters are used to determine the color. The scale parameter (resembling the average wind speed) defines the brightness, while the shape parameter defines the hue. Blue areas have a shape parameter close to 1, meaning broad distributions, while red areas show narrower distributions, which are more centered around the average. Here, we also observe a good spatial continuity in the fitted distributions. The large valleys all have a low scale parameter, which increases towards the summits and ridges. However, in many locations the highest values are not found there, but slightly below them. Also, the concave areas on slopes exhibit higher scale parameters than their surroundings.

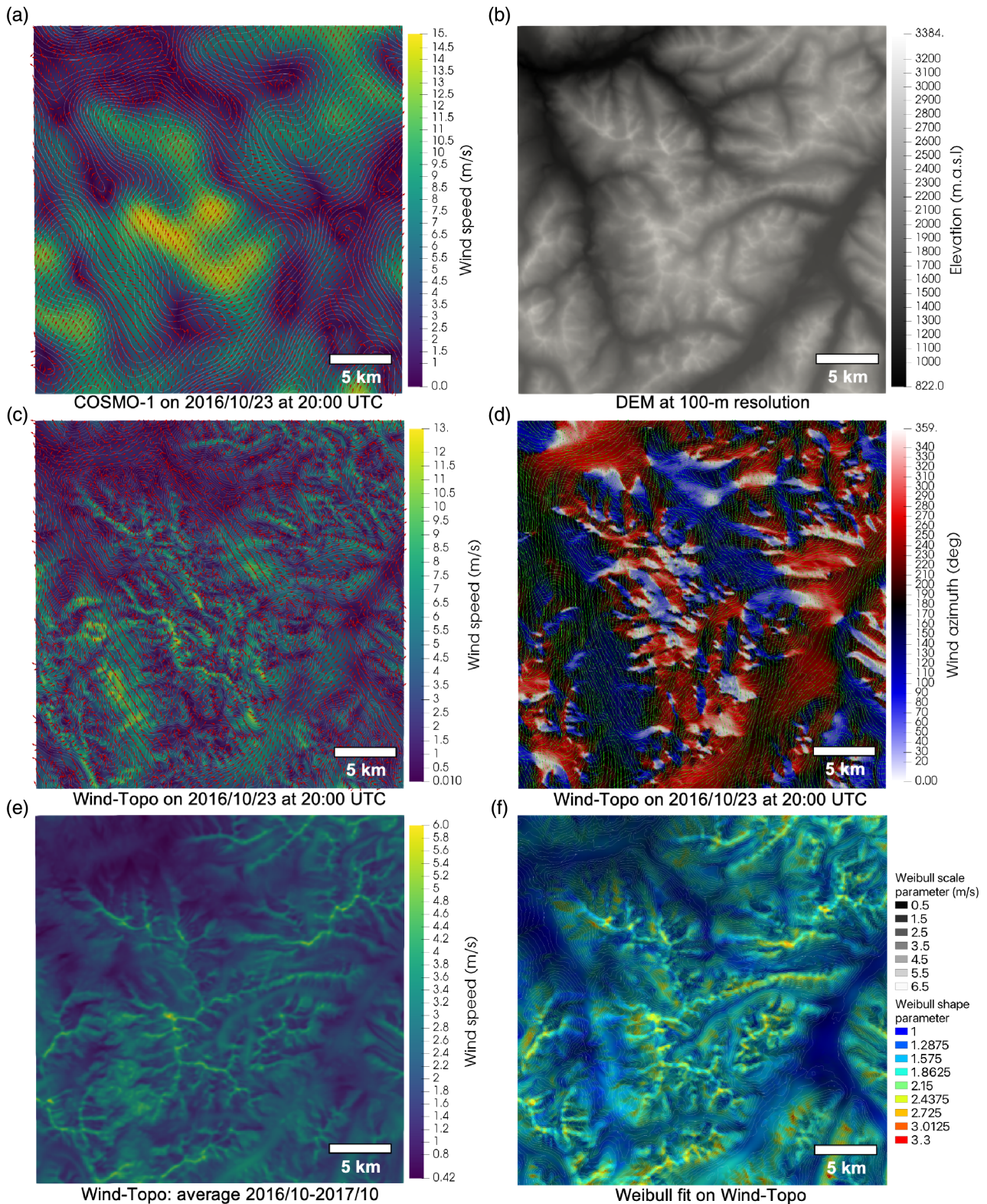


FIGURE 12 (a) Wind field from COSMO-1, interpolated (bicubic) on a 100-m grid, for a specific time. The background color indicates the wind speed and red arrows indicate the wind direction every 500 m. Grey lines show the 100-m contours of the associated topography. (b) 100-m digital elevation model used by Wind-Topo covering the same domain. (c) 100-m wind field generated by Wind-Topo for the same time. Grey lines show the underlying topography (100-m contours). (d) Associated wind direction (clockwise from North) with a cyclic color scale. The green lines are the stream lines of the wind field, with arrows indicating the flow direction. (e) Average wind speed from Wind-Topo over the test period. (f) Weibull parameters of the distribution fitted in each pixel. The brightness indicates the scale parameter and the color indicates the shape parameter. 100-m contours in grey show the underlying topography [Colour figure can be viewed at wileyonlinelibrary.com]

A similar plot for COSMO-1 is provided in Appendix S1 and shows the same increase in concave areas on larger scales. This local increase can be explained by the channeling that occurs in such places, which generates many situations with relatively high wind speed but few occurrences of a very large one. Wind-Topo is able to reproduce such an effect on smaller scales.

One might wonder about the physical consistency of the wind fields generated. As Wind-Topo only predicts a near-surface wind field, we cannot compute a 3D divergence to assess the mass conservation. We can nevertheless estimate the ground-perpendicular motion caused by the divergence of the 2D field. Appendix S1 describes this analysis along with figures, which we summarize here. The large majority of such motions at 7 m a.g.l. are between -0.04 and $0.04 \text{ m}\cdot\text{s}^{-1}$ (negative for motions towards the ground). Rarely, those motions can be 10 times larger. When averaged for the entire year, we obtain values between -0.06 and $0.05 \text{ m}\cdot\text{s}^{-1}$. Furthermore, if we apply a spatial averaging to coarse-grain these maps to the scales resolved by COSMO-1, we obtain patterns that are very similar to those of COSMO-1 in terms of magnitude (-0.01 to $0.006 \text{ m}\cdot\text{s}^{-1}$) and location.

4 | CONCLUSION

A high spatial resolution in complex terrain is essential to capture small-scale processes shaping the wind flow that is responsible for phenomena like snow preferential deposition or transport. Even state-of-the-art operational models like COSMO-1 running on supercomputers cannot provide the required resolution for topographies like the Alps. Wind-Topo offers the possibility to extract the most information from such models and to correct the near-surface wind fields to match observations closely. The systematic biases observed in different types of terrain and at different times of the year and day are strongly reduced. The downscaled wind fields exhibit some of the expected orographic effects (ridge acceleration, sheltering, and deflection), and Wind-Topo does not alter the predictions from the low-resolution input model in areas where it already performs well. At 44 measurement stations located in diverse topographies in the Alps and put aside for validation, Wind-Topo offers an average correlation of 0.55, a low bias of $-0.07 \text{ m}\cdot\text{s}^{-1}$, and a MAE of $1.21 \text{ m}\cdot\text{s}^{-1}$. This should be compared with the values obtained by a simple interpolation of COSMO-1: correlation of 0.50, bias of $0.72 \text{ m}\cdot\text{s}^{-1}$, and MAE of $1.77 \text{ m}\cdot\text{s}^{-1}$.

With its novel architecture, Wind-Topo not only downscales the wind speed to higher resolutions but also incorporates local orographic deflections. Even if subgrid-scale thermally driven flows are not captured currently, the level

of performance is remarkable and further improvements are already foreseen, in particular if more diverse locations with observed or modeled “ground truth” can be used for training. Given the rapid evolution of the research on deep learning, we believe that its architecture will also be improved in the near future.

The point-by-point nature of Wind-Topo can be seen as a strength or as a weakness. Using point observations as predictands avoids the introduction of inaccuracies or biases that other (modeled) data sources would have. The pointwise prediction also simplifies the parallelization of the downscaling scheme. However, it prevents the introduction of physical constraints during the training and predictions that would guarantee the generation of physically consistent flows. Ideally, Wind-Topo should be developed further to incorporate high-resolution 2D or 3D modeled data as well, which could help identify wind-topography interactions not yet discovered by the model. This new data source could also provide the physical constraints needed when generating wind fields.

Wind-Topo is already quite fast and can be optimized further. In our Swiss case study, it can downscale COSMO-1 (1,110 m) to 6 million grid points in one hour on one standard GPU. It is thus possible to downscale the Swiss Alps (about 29,000 km²) operationally (hourly) to a 50-m grid with a 2-GPU machine. As the computational requirements scale linearly with the number of grid points desired, Wind-Topo can be applied to much larger domains at this resolution with reasonable resources. In Appendix S1, we compare the current version of the model with simpler models and lighter model variants. The 2D nature of the model provided by its CNNs is of high importance. The custom loss function is also crucial in obtaining correct distributions of wind speed.

We hope this work will pave the way for new methods based on deep learning that will downscale not only the wind but also other atmospheric variables to higher resolutions in complex terrain. Such methods will be able to combine state-of-the-art high-resolution NWP models, observations, and operational models efficiently.

AUTHOR CONTRIBUTIONS

Jérôme Dujardin: conceptualization; investigation; data collection; curating; development; validation; writing - original draft; writing - review and editing. **Michael Lehning:** conceptualization; funding acquisition; investigation; supervision; writing - original draft; writing - review and editing.

ACKNOWLEDGEMENTS


We thank Mathieu Schaer for the first investigations on Wind-Topo during his master's project and Quentin Fisch for his help in parallelizing the code. This work was funded

by Innosuisse through the Swiss Competence Center for Energy Research, Supply of Electricity.

DATA AVAILABILITY STATEMENT

The original datasets used to train and validate Wind-Topo can be obtained free of charge from the data owners MeteoSwiss and the WSL Institute for Snow and Avalanche Research. Intermediate data products can be obtained from the authors upon request. A complete use-case dataset is available from the ENVIDAT data platform under the DOI: 10.16904/envidat.301. It includes the architecture of Wind-Topo and its optimized parameters, as well as a PYTHON script to downscale uniform wind fields with a prescribed vertical profile for any given 53-m DEM. New versions of the code will be available on the WSL GitLab: <https://gitlabext.wsl.ch/dujardin/wind-topo>.

ORCID

Jérôme Dujardin  <https://orcid.org/0000-0001-5404-7734>

REFERENCES

- Burger, W. and Burge, M.J. (2009) *Principles of Digital Image Processing: Core Algorithms*. London: Springer.
- Dupuy, F., Duine, G.J., Durand, P., Hedde, T., Pardyjak, E. and Roubin, P. (2021) Valley winds at the local scale: correcting routine weather forecast using artificial neural networks. *Atmosphere*, 12, 28.
- Etienne, C., Lehmann, A., Goyette, S., Lopez-Moreno, J.-I. and Beniston, M. (2010) Spatial predictions of extreme wind speeds over Switzerland using generalized additive models. *Journal of Applied Meteorology and Climatology*, 49, 1956–1970.
- Fiddes, J. and Gruber, S. (2014) TopoSCALE v.1.0: downscaling gridded climate data in complex terrain. *Geoscientific Model Development*, 7, 387–405.
- Fischer, P., Etienne, C., Tian, J. and Krauß, T. (2015) Prediction of wind speeds based on digital elevation models using boosted regression trees. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1-W5, 197–202. <https://elib.dlr.de/99861/1/FischerEtAl.pdf>
- Foresti, L., Tuia, D., Kanevski, M. and Pozdnoukhov, A. (2011) Learning wind fields with multiple kernels. *Stochastic Environmental Research and Risk Assessment*, 25, 51–66.
- Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* Vol. 9, pp. 249–256. <https://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf>
- Goutham, N., Alonzo, B., Dupré, A., Plougonven, R., Doctors, R., Liao, L., Mougeot, M., Fischer, A. and Drobinski, P. (2021) Using machine-learning methods to improve surface wind speed from the outputs of a numerical weather prediction model. *Boundary-Layer Meteorology*, 179, 133–161. <https://doi.org/10.1007/s10546-020-00586-x>.
- Helbig, N., Mott, R., Van Herwijnen, A., Winstral, A. and Jonas, T. (2017) Parameterizing surface wind speed over complex topography. *Journal of Geophysical Research*, 122, 651–667. <https://doi.org/10.1002/2016JD025593>.
- Hendrycks, D. and Gimpel, K. (2016) *Gaussian Error Linear Units (GELUs)*. <https://doi.org/10.48550/arXiv.1606.08415>.
- Hilton, J. and Garg, N. (2021) Rapid wind-terrain correction for wildfire simulations. *International Journal of Wildland Fire*, 30, 410–427.
- Höhlein, K., Kern, M., Hewson, T. and Westermann, R. (2020) A comparative study of convolutional neural network models for wind field downscaling. *Meteorological Applications*, 27, 1–31.
- Hsieh, W.W. (2009) *Machine Learning Methods in the Environmental Sciences: Neural Network and Kernels*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511627217>
- Huang, H.Y., Capps, S.B., Huang, S.C. and Hall, A. (2015) Downscaling near-surface wind over complex terrain using a physically-based statistical modeling approach. *Climate Dynamics*, 44, 529–542.
- Kingma, D.P. and Ba, J.L. (2015). ADAM: a method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015*, pp. 1–15. <https://arxiv.org/pdf/1412.6980.pdf>
- Kruyt, B., Dujardin, J. and Lehning, M. (2018) Improvement of Wind Power Assessment in Complex Terrain: The Case of COSMO-1 in the Swiss Alps. *Frontiers in Energy Research*, 6, 102. <https://doi.org/10.3389/fenrg.2018.00102>.
- Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U. and Zimmerli, M. (1999) SNOWPACK model calculations for avalanche warning based upon a new network of weather and snow stations. *Cold Regions Science and Technology*, 30, 145–157.
- Manor, A. and Berkovic, S. (2015) Bayesian inference aided analog downscaling for near-surface winds in complex terrain. *Atmospheric Research*, 164–165, 27–36. <https://doi.org/10.1016/j.atmosres.2015.04.014>.
- Mott, R., Schirmer, M., Bavay, M., Grünewald, T. and Lehning, M. (2010) Understanding snow-transport processes shaping the mountain snow-cover. *Cryosphere*, 4, 545–559.
- Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L. and Lerer, A. (2017). Automatic differentiation in PyTorch. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. <https://openreview.net/pdf?id=BJJrmfCZ>
- Philippopoulos, K. and Deligiorgi, D. (2012) Application of artificial neural networks for the spatial estimation of wind speed in a coastal region with complex topography. *Renewable Energy*, 38, 75–82. <https://doi.org/10.1016/j.renene.2011.07.007>.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N. and Prabhat (2019) Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195–204. <https://doi.org/10.1038/s41586-019-0912-1>.
- Rios, O., Jahn, W., Pastor, E., Valero, M.M. and Planas, E. (2018) Interpolation framework to speed up near-surface wind simulations for data-driven wildfire applications. *International Journal of Wildland Fire*, 27, 257–270.

- Robert, S., Foresti, L. and Kanevski, M. (2013) Spatial prediction of monthly wind speeds in complex terrain with adaptive general regression neural networks. *International Journal of Climatology*, 33, 1793–1804. <https://doi.org/10.1002/joc.3550>.
- Takeuchi, I., Le, V.Q., Sears, T. and Smola, A. (2006) Non-parametric Quantile Estimation. *Journal of Machine Learning Research*, 7, 1231–1264. <https://jmlr.org/papers/volume7/takeuchi06a/takeuchi06a.pdf>.
- Vionnet, V., Marsh, C.B., Menounos, B., Gascoin, S., Wayand, N.E., Shea, J., Mukherjee, K. and Pomeroy, J.W. (2021) Multi-scale snowdrift-permitting modelling of mountain snowpack. *Cryosphere*, 15, 743–769.
- Wagenbrenner, N.S., Forthofer, J.M., Lamb, B.K., Shannon, K.S. and Butler, B.W. (2016) Downscaling surface wind predictions from numerical weather prediction models in complex terrain with WindNinja. *Atmospheric Chemistry and Physics*, 16, 5229–5241.
- Winstral, A., Elder, K. and Davis, R.E. (2002) Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of Hydrometeorology*, 3, 524–538.

- Winstral, A., Jonas, T. and Helbig, N. (2017) Statistical downscaling of gridded wind speed data using local topography. *Journal of Hydrometeorology*, 18, 335–348. <https://doi.org/10.1175/JHM-D-16-0054.1>.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Dujardin, J. & Lehning, M. (2022) Wind-Topo: Downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning. *Quarterly Journal of the Royal Meteorological Society*, 148(744), 1368–1388. Available from: <https://doi.org/10.1002/qj.4265>