Wind resource assessment using a novel meso-microscale coupling framework based on two-level self-organized maps clustering

Pablo Durán [a, *], Arne Gravdahl [a]

[a] WindSim AS, Fjordgaten 13B, Tønsberg, Norway

* Corresponding author. E-mail address: pablo.duran@windsim.com

Abstract

A novel meso-microscale coupling framework is proposed which uses a fully automated classification approach to extract few predominant classes from a mesoscale simulation. The mesoscale classes are physically downscaled using a steady-state CFD model by transferring the mesoscale information as boundary conditions of the CFD model. We apply the classification to a site located in a desert area and demonstrate its capability to organize the extracted classes according to atmospheric stability, wind speed and wind direction and how they relate to seasons and diurnal cycles. The downscaled classes are validated against onsite measurements and compared to standalone CFD simulation. We show that the meso-microscale coupling simulations reduce the horizontal extrapolation error of the resource in 3.24% compared with the standalone simulation.

Keywords: RANS, WRF, CFD, Physical downscaling, Diurnal cycle, Mesoscale models, clustering.

1. Introduction

Current evaluation of wind energy projects heavily relies on the use of steady-state wind flow models to estimate the energy production. In order to utilize steady-state models in wind resource assessments, it is widely assumed in that the ratio of the horizontal wind speed between two points is roughly constant for a given wind direction. This ratio, referred hereafter to as speed-up (SU), is utilized to extrapolate wind speeds from a reference point u_R (usually measurements) to a target point u_T (usually wind turbines) by using the relation $SU = \frac{u_T}{u_R}$. The assumption that SU values are time-independent and only dependent on the wind direction allows to lower the use of computational resources without losing significant accuracy in the estimations. However, for some sites this assumption does not hold true, like in the example case shown in Figure 1. For this site, a unique SU value for the direction 0°±15° may be reasonable for relatively high wind speeds. Nevertheless, at lower wind speeds (but still well above 3 m/s), this is not the case. Furthermore, at this range of lower speeds, there might be

non-linear dependences between the observed SU values and the observed atmospheric stability (accounted by the temperature gradient).

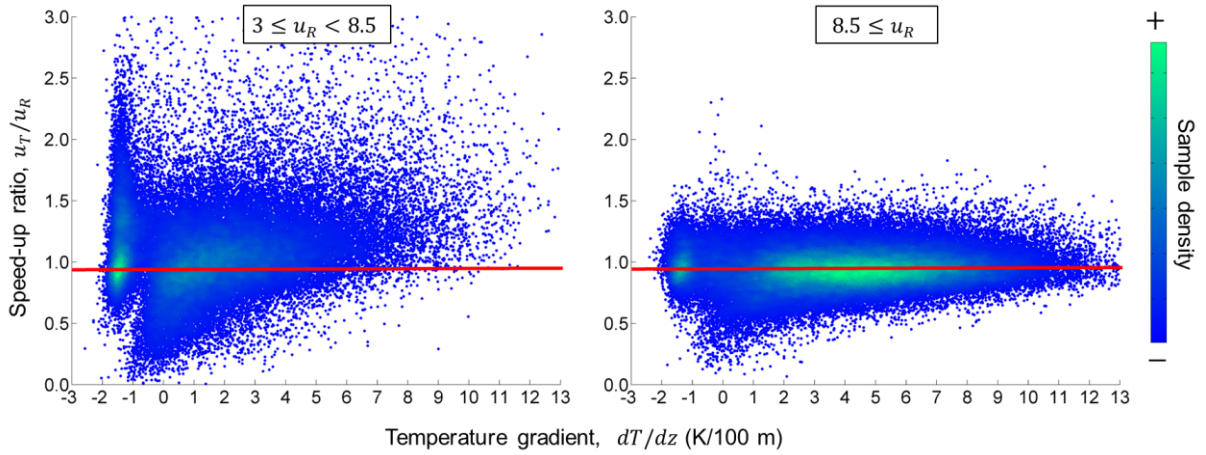

Figure 1. Example of measured speed-up ratios for the directional sector 0°±15°. The considered values of u_R (m/s) are indicated at the top of each panel. The red line indicates a $SU = 0.95$.

For the example of Figure 1, it can be expected that a steady-state model would calculate a SU value of ~0.95. Such a value would be not appropriate for relatively low wind speeds. A possible approach to deal with this limitation would be to simulate different atmospheric stability conditions for the same directional sector. Even assuming the minimum possible number of atmospheric stability classes (unstable, neutral, and stable), the total number of simulations would triple. Furthermore, some combinations of wind direction and stability will probably have very low or no occurrence at the modelled site, wasting computational resources. Another alternative to obtain SU values that vary with u_R is to use a transient model instead of a steady-state one, avoiding the assumption of constant SU values per direction at all. However, this type of computational fluid dynamic (CFD) models is regarded as computationally expensive for most of the industry.

To deal with the kind of sites of Figure 1, we propose a wind resource assessment framework based on the coupling of a mesoscale and a microscale model. An important characteristic of the framework is that instead of utilizing single SU values per directional sector, it uses one SU value per predominant wind condition at the site. These "wind conditions" correspond to a combination of local wind speed, wind veer, wind shear and atmospheric stability and are obtained from the mesoscale simulation. The most predominant site-specific wind conditions

are extracted in an automated manner utilizing a machine learning-based approach and then transferred to the CFD model as boundary conditions. The microscale CFD model resolves the wind flow using these boundary conditions, obtaining a set of SU values per downscaled condition.

Details about the studied site and its corresponding datasets are presented in Section 2. Technical details about the CFD model, mesoscale classification methodology, coupling procedure and validation methodology are presented in Section 3. In Section 4, we discuss the mesoscale patterns obtained by the automatic classification. In addition, in the same section validation results of the coupled and standalone CFD simulations are shown. Further discussion and conclusions of this exercise are presented in Section 5.

2. Validation site and datasets

The validation study is carried out using a site hereafter referred to as CK. This site belongs to a commercial wind energy generation project and therefore information regarding the real name, exact location, absolute elevation, and absolute wind speeds are not disclosed. The site is located in the Southern Cone of South America, in a desert area. Two meteorological masts are located in the area, M1 and M2, and one LIDAR referred to as L1. The wind conditions at the site are strongly influenced by the diurnal cycle, with unstable atmospheric conditions during the 1100 and 1800 LST and stable atmospheric conditions during 2200 and 1000 LST (Figure 2). A summary of the terrain and onsite measurement data is presented in Figure 3. Further details of the datasets of the site are presented in the following subsections.

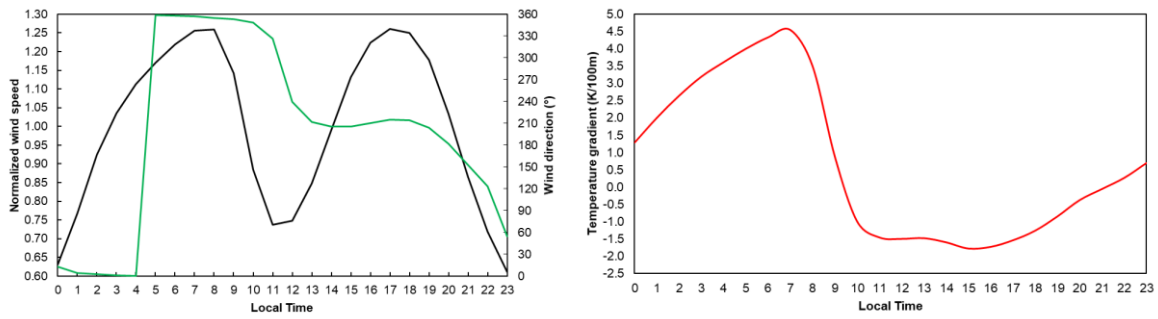

Figure 2. Left-panel: Measured diurnal profiles of wind speed (black) and direction (green) at M2@80m. Right-panel: Measured diurnal profiles of temperature gradient between 80 and 15 m at M2.

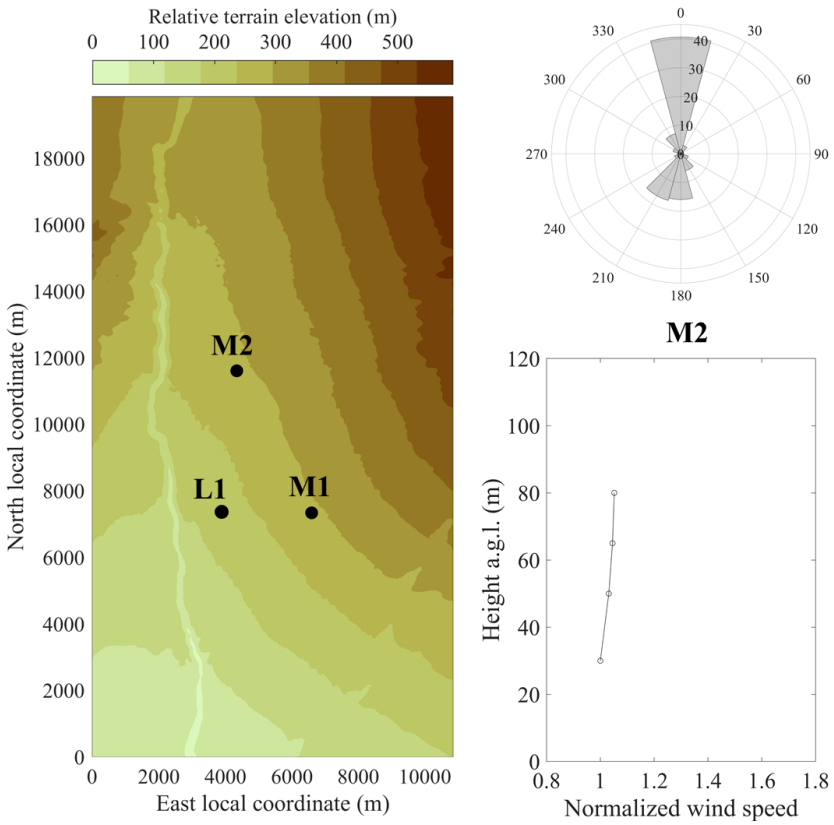

Figure 3. Left-panel: terrain elevation map of the CK site. The location of the measurement points is indicated by black circles. Right panels: wind rose @78m (top) and vertical profile of wind speed (bottom). The instrument used to compute them is indicated above the bottom-right panel.

2.1. Terrain data

The terrain data consist of the map of terrain elevation and roughness. The elevation dataset was retrieved from the Shuttle Radar Topography Mission (SRTM) database.[1] This database also contains land cover data, from which roughness can be derived. Nevertheless, the quality of the data was found to be of poor quality for the studied area and therefore a constant roughness length =0.01 m was used instead as the terrain consist of mostly sand. The site is relatively flat and with a positive inclination northward (Figure 3). A riven is located at the west part of the site that traverses from south to north.

2.2. Wind measurement data

Wind conditions at the site were monitored through anemometers and wind vanes mounted in the two meteorological masts and a LIDAR. Ten-minute averages of horizontal wind speed and wind direction at several heights are available (Table 1). The location of each measurement is shown in Figure 3. As shown in the same figure, the wind at the site has on average a very low shear and most of the wind comes from the north. Each of the two meteorological masts have more than 7 years of coverage, while the LIDAR covers 5 months.

Table 1. Details of the onsite measurements at CK.

Measurement name	Type of instrument	Measured Heights a.g.l. (m)	Recorded period (month/year)
M1	Anemometer	10, 20, 30, 40, 50, 60	04/2010 – 10/2018
	Wind vane	28, 48, 78	04/2010 – 10/2018
M2	Anemometer	30, 50 ,65, 80	10/2011 – 10/2018
	Wind vane	28, 63, 78	10/2011 – 10/2018
L1	LIDAR	18, 28, 38, 78, 98, 128	06/2010 – 11/2010

2.3. Mesoscale simulations

One-year worth of mesoscale simulations are produced using the Weather research and Forecasting (WRF) model. The simulation period covers from October 2013 to October 2014. This timespan was selected in order to cover the longest concurrent measured year among the measurements presented in Table 1 and others not used in this study due to limited coverage. The configuration of the WRF model, including physics parameterizations and grid settings are the same as the ones used in Durán et al. (2020),[2] where the same site and datasets are also studied. The innermost domain has a resolution of 1 km and it was configured to cover the whole microscale domain. Mesoscale outputs are stored every 1 hr and therefore a total of 8760 timesteps are available.

3. Methodologies

In this work, an automated classification methodology is utilized to extract the dominant patterns from the mesoscale simulation. The obtained mesoscale fields are then used to compute the boundary conditions of a steady-state microscale model. Details about the microscale model, the classification methodology and the procedure to compute the boundary

conditions are provided in the following subsections. Finally, the validation procedure is presented.

3.1. Microscale steady-state CFD model

The CFD software WindSim is used to simulate the microscale wind flow. WindSim's governing equations are based on the steady-state version of the Reynolds-averaged Navier-Stokes (RANS) equations. As turbulence closure scheme, the standard k-ε model[3] available in the software is used. For more details about the actual formulation of the governing equations in WindSim, the reader is referred to the work of Gravdahl (1998).[4] Thermal effects are considered by introducing a temperature equation and extra source/sink terms to the momentum and turbulence equations to represent buoyancy, as detailed by Meissner et al (2009).[5] For a given set of boundary conditions, the microscale model provides a wind flow solution. The common practice in the wind industry is to use a set of 12 or 16 boundary conditions, each with a different wind direction, typically evenly spaced.

The CFD digital terrain model used in this study covers a rectangular grid of 11 x 20 km at the CK site. The terrain model has a concentric refinement area of 5.3 x 13.6 km with a horizontal resolution of 40 m. The three measurement locations lie within the refinement area. From the boundaries of the refinement area towards the boundary of the domain, the horizontal resolution is gradually reduced up to 170 m. A total of 60 vertical levels is used, spanning up to 4000 m a.g.l. The vertical levels are distributed following an arithmetic progression with the highest resolution towards the ground. 9 vertical levels are located within the first 100 m.

3.2. Automated classification of simulated mesoscale fields

An automated classification methodology is used to extract the most predominant mesoscale fields from the 3-D timeseries. The classification methodology consist of a clustering technique that uses a two-level self-organizing map clustering (SOM2L) approach.[6] This SOM2L allows for a fully automated clustering of the data without the need to prescribe the number of clusters a priori, making it more advantageous than other popular clustering techniques like *k*-means. The first level or stage of the SOM2L corresponds to the training of a self-organizing map (SOM)[7] with the input data. The SOM consists of interconnected nodes (or neurons) in a 2-D grid array that are iteratively adapted according to the input data. After the training, the

SOM provides a visualization of the clustering structures by plotting the distances between the nodes. The second level of the SOM2L algorithm exploits this feature to determine the number of clusters within the data.[8] Further technical details of the application of the SOM2L are presented in the following subsections.

SOM2L input data

As input data, only the mesoscale model grid points located within the microscale domain are considered, corresponding to a total of 11 x 21 x 19 = 4389 points. In addition, the mesoscale data is filtered by removing the timesteps in which the average wind speed between 50 and 150 m is lower than 3 m/s. This way, only timesteps that are of interest for wind energy generation are considered in the classification. The total number of samples after the filtering is 7740. The considered variables for the classification are the wind velocity components in the east direction U_x and in the north direction U_y, as well as the gradient of potential temperature $\Delta\theta = \theta - \theta_0$. Here, θ and θ_0 is the potential temperature at all heights and at the lowest vertical level of the WRF model, respectively. U_x and U_y fields contain information about the wind direction, wind veer and wind shear, while $\Delta\theta$ contains information about the atmospheric stability. In summary, the input data used in the classification consist of 7740 samples, each corresponding to a vector of length =12936 (=[11 x 21 x 19] +[11 x 21 x 19] + [11 x 21 x 18]).

SOM2L level 1 – SOM training

A 25 x 25 SOM with hexagonal grid cells is used, giving a total of 625 nodes. This is much higher than the value given by the rule of thumb proposed by Vesanto et al. (2000),[9] which suggests the use of approximately 468 nodes for a sample size of 8760. This rule of thumb works as a lower limit since the SOM should have as many nodes as possible when used for clustering purposes, as pointed out by Kohonen (2013).[10] The SOM is trained with the input data: First, the SOM is initialized using linear initialization[7] and then trained using the batch algorithm[7] in two training phases. The first phase uses a neighborhood radius =25 that linearly decreases to 1 in 200 iterations. The second phase uses a radius =1 for another 200 iterations. The trained SOM allows to identify the clustering structures within the dataset (Figure 4a). In this figure lighter (darker) coloring represents samples in the dataset that are more similar (different) to each other.

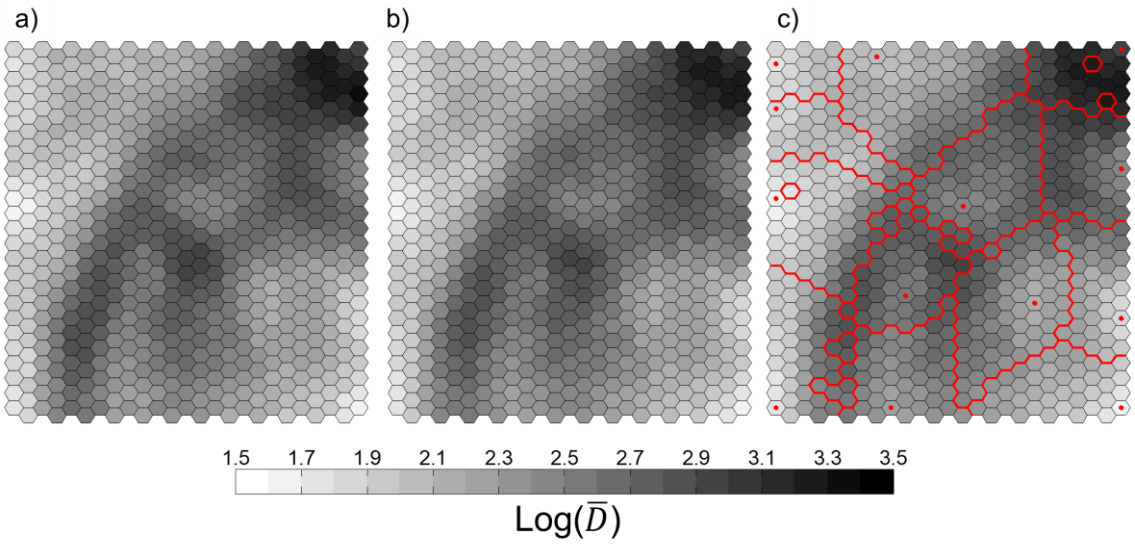

Figure 4. a) Distance map of the trained SOM applied to the CK mesoscale data, b) distance map after the smoothening and c) partition of the SOM based on the local minima of $\bar{D}$. In c) the local minima are indicated by the red dots and the partitions by the red lines. $\bar{D}$ corresponds to the average distance between a node and its neighbors.

SOM2L level 2 – SOM partitioning

After, the SOM is trained, the map is partitioned to identify the classes. The partition of the SOM is based on the distances between the nodes (Figure 4). These distances are roughly inversely proportional to the probability density function of the input data.[8] Therefore, it is expected that the local minima of the distances correspond to the cluster centers of the data. A potential problem with this approach is that the distances could have some local minima which are a product of random variations in the data, rather than actual local minima.[8] For this reason we apply a fully automated smoothening procedure proposed by Garcia (2010)[11] before computing the local minima, resulting in the map in Figure 4b. After the smoothening, initial clusters are defined as the nodes corresponding to the local minima (Figure 4c). Neighboring nodes of these initial clusters are assigned to a cluster based on Ward's criteria,[12] which minimizes the variance within the cluster. Some of the enclosed areas in Figure 4c do not have a local minima. The reason is that the nodes in these areas do not have assigned samples. This kind of nodes are referred in the literature as "interpolating nodes" and represent the inter-cluster borders on the SOM. Therefore, they are used to delimit the classes and are not assigned to a class.

3.3. Microscale boundary conditions based on mesoscale classes

For each of the mesoscale classes obtained by the SOM2L, one mesoscale field is computed by averaging all the members belonging to the same class. Following Duran et al (2020),[2] U_x and U_y are not directly averaged and instead $u\left(=\sqrt{U_x^2+U_y^2}\right)$ and $\beta\left(=\arctan\frac{U_y}{U_x}\right)$ are separately averaged to then obtain the resulting components. In the case of the potential temperature, $\Delta\theta$ and θ_0 are averaged separately to then obtain the values of θ of the mesoscale field. The other mesoscale variables used in the coupling procedure, the vertical component of the wind U_z and the planetary boundary layer height h, are directly averaged. This procedure yields a single mesoscale field per class.

Each of the resulting mesoscale fields are transferred to the microscale model as initial and boundary conditions. In a nutshell, the wind and potential temperature fields are interpolated into the microscale domain. For microscale grid points below the lowest vertical level of the mesoscale model, the values are extrapolated using the Monin-Obukhov similarity equations. The turbulent kinetic energy and its energy dissipation rate are prescribed in the microscale model using the analytical formulations derived by Han et al. (2000),[13] but using as input parameters values obtained from the mesoscale. For further details of the transferring procedure, the reader is referred to the work of Durán et al. (2019).[14]

3.4. Validation procedure

As mentioned in the introduction, when using CFD models for wind resource assessment the quantity of interest are the SUs values. Since one SU value per simulated wind direction is obtained, the measurements used for validation are filtered according to the measured wind direction. However, for the proposed framework of this study, the measured wind direction is not sufficient to filter the measurements, since it is very probable that 2 or more classes are obtained for a given wind direction (specially for prevailing directions). The proposal in this study is to assign a given measured timestep to the selected class k^*, such as:

$$k^* = \arg\min_{k\in K}\left(\sum_{p\in P}\left|U_{x,k}(p)-U_{x,m}(p)\right|+\left|U_{y,k}(p)-U_{y,m}(p)\right|\right) \quad (1)$$

Here, $U_{xk}(p)$ and $U_{yk}(p)$ correspond to the east and north components of the wind simulated for the class k at the observational point p, respectively. On the other hand, $U_{x,m}(p)$ and $U_{y,m}(p)$ correspond to east and north components of the measured wind at the observational point p, respectively. The set K consists of all the classes obtained by the SOM2L method and the set P consists of the observational points considered for the assignment. In a nutshell, equation (1) indicates that a given measured timestep is assigned to the class k^* by comparing the measured horizontal components of the wind with the simulated ones. A class for the timestep is selected by minimizing the difference. In this study, P corresponds to the measurement points at M1 @ 50 and 60 m, and M2 @ 50, 65 and 80 m.

After each timestep is assigned to a class, the simulated speed-up for the class k between the reference point R and the target point T, denoted as $SU_k(R,T)$, is validated. The common practice in the industry is to indirectly validate it by comparing the predicted wind speed using a given SU and the measured wind speed. The error of this prediction is referred to as the crosscheck prediction error (XPE) and is calculated as:

$$XPE_k(R,T) = \frac{SU_k(R,T) \times \overline{u_{Rk}} - \overline{u_{Tk}}}{\overline{u_{Tk}}} \times 100\% \quad (2)$$

Here $\overline{u_{Rk}}$ and $\overline{u_{Tk}}$ are the average measured wind speed at points R and T for the class k. One value of $XPE_k(R,T)$ is obtained per pair of measurements. In order to facilitate the error analysis, the mean absolute average of the XPE values per class are calculated as:

$$AXPE_k = \frac{\sum_{R,T \in M; R \neq T} |XPE_k(R,T)|}{N_M} \quad (3)$$

Here M is the set of the observational points considered for the error calculation and N_M is the total number of pair of observational points in M.

4. Results

Two types of results are discussed in the following subsections. First, the extracted mesoscale fields obtained by the SOM2L are presented and characterized. Secondly, the downscaled classes are validated against onsite measurements and compared with standalone microscale simulations.

4.1. Application of SOM2L to the mesoscale timeseries

This subsection illustrates the capabilities of the SOM2L method to facilitate the identification and understanding of the prevalent wind conditions at a site. The SOM2L is applied to the 3-D mesoscale fields generated with the WRF model for the CK site, obtaining a total of 13 classes. One important advantage of using a SOM in the classification procedure is that it provides spatial ordering of the underlaying input data. This means that areas that are adjacent to one another in the SOM share certain characteristics (in this case wind speed, wind shear, wind direction, wind veer and/or $\Delta\theta$), and therefore the same is true for adjacent classes. This feature of the SOM facilitates the characterization of the obtained cluster as well as their relation (Figure 5). It has to be underlined that this ordering is obtained from a fully automated procedure.

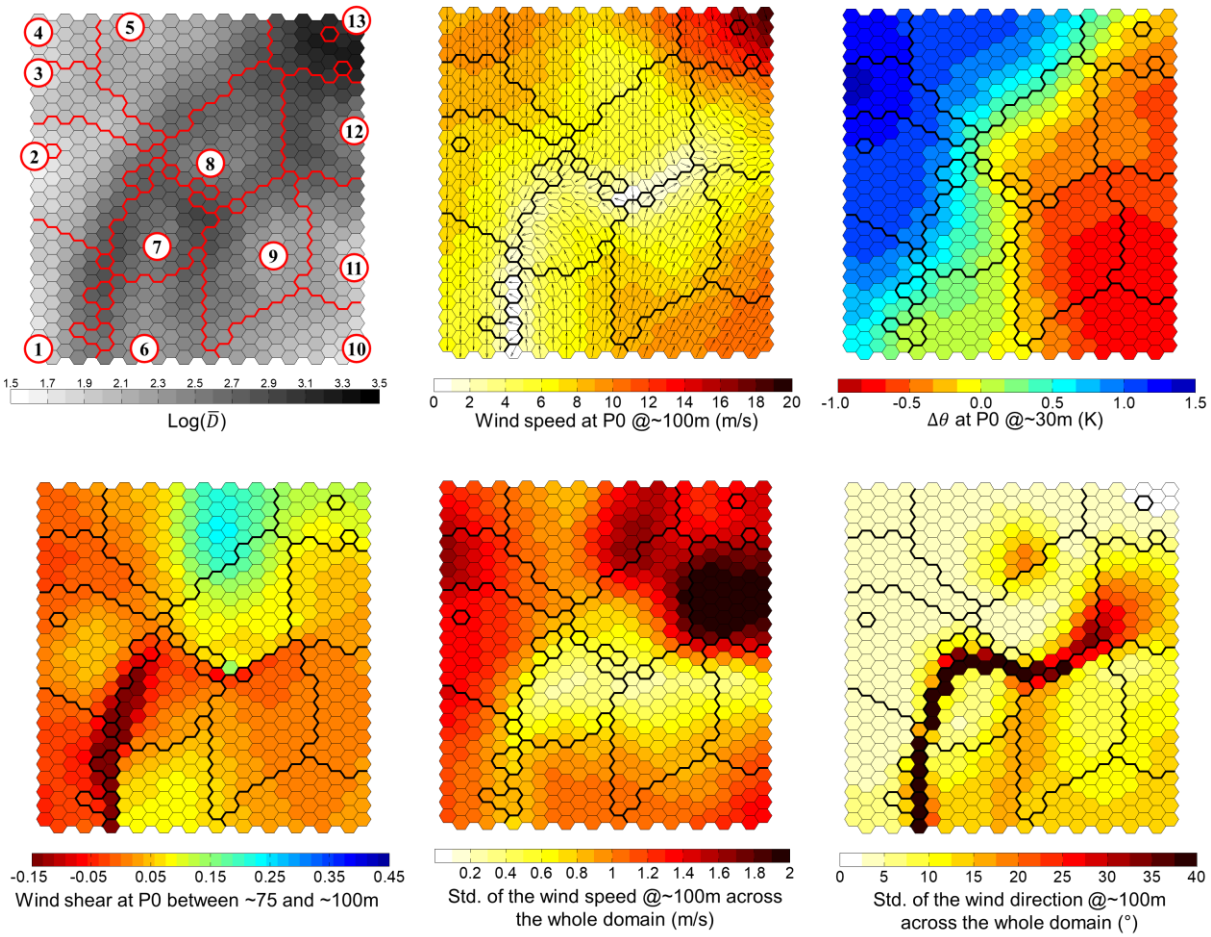

Figure 5. Top left panel: Partition of the SOM with the class number indicated. Rest of the panels: Representative values of different variables in the nodes of the SOM. The bold lines between the nodes

indicate the boundaries between the classes. The black arrows in the top-center panel correspond to the wind direction. P0 correspond to the center point of the domain.

The panels in Figure 5 provide additional information on how the classes relate to each other. While the $\overline{D}$ can provide a more abstract but general understanding on how similar/dissimilar the adjacent nodes of the SOM are, the other presented variables help us to understand this relation in more familiar quantities. For example, most classes have either wind coming from the north (classes 1, 2, 3, 4, 5, 8 and 13) or from the south (classes 6, 9, 10 and 11). Furthermore, most of the classes with northern wind are correlated with stable atmospheric conditions, whereas southern wind classes are correlated with unstable conditions. There are also some classes that exhibit interesting patterns. For example, Class 3 in addition to have stable stratification and north winds, it also presents negative shear, which indicates that corresponds to a low-level jet, as confirmed in Figure 6. Class 7 has a relatively low standard deviation of the wind speed, which is due to a relatively homogeneous wind field, as seen in Figure 7. On the other hand, class 10 has a relatively high standard deviation of the wind direction, which can be seen in Figure 7 that is due to a considerable horizontal wind veering across the domain.

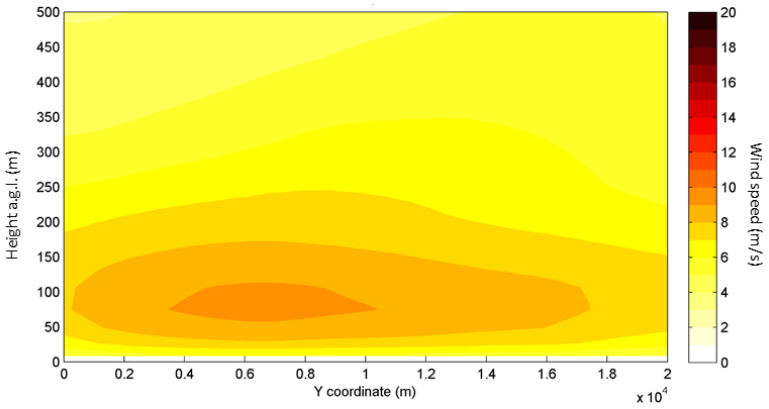

Figure 6. Vertical plane of horizontal wind speed of the mesoscale data of class 1. The vertical cut goes from south to north through the center of the domain.

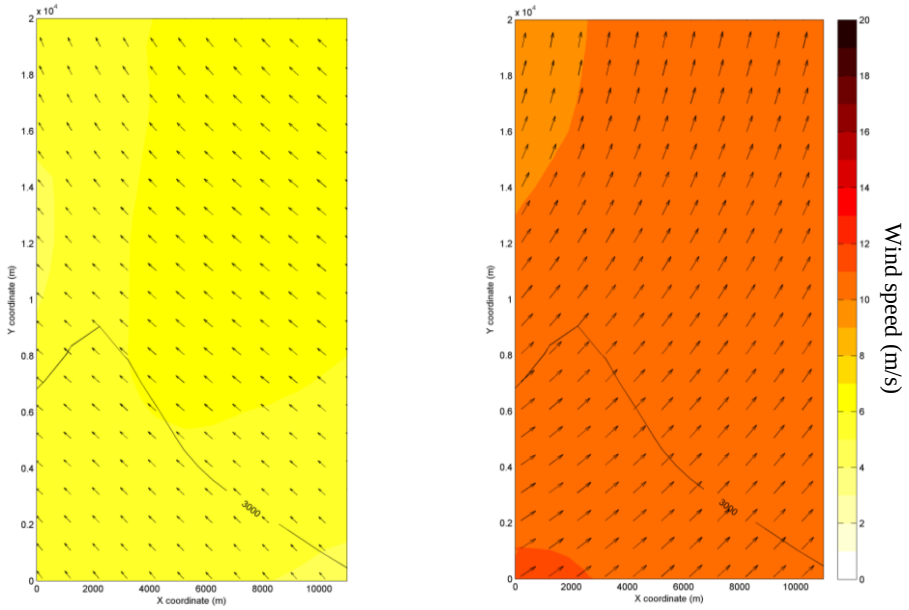

Figure 7. Horizontal plane @~100m of horizontal wind speed of the mesoscale data of classes 7 (left panel) and 10 (right panel).

In Figure 8, the most frequent class for a particular combination of month and time of the day is shown. This information allows us to characterize temporally the classes, as well as connecting the site conditions they represent, like atmospheric stability, to certain moments of the year of day. For example, northern wind classes like 1, 2, 3, 5 and 8 tend to occur between 00:00 and 09:00. After that, between October and April, the wind direction changes to south (classes 6, 9 and 10), which persist for the rest of the day. Between May and September, the wind changes to a northeast direction during daytime (class 12) and then transitions through class 7 with southwest winds, to finally recover to northern winds after 22:00. The time of occurrence and negative shear of the norther classes indicates that they most likely correspond to nocturnal low-level jets. On the other hand, it must be noted that the northern wind classes represent different wind speed intensities and tend to occur at different months. This information allows to determine that northern winds in summer are stronger (like class 5) compared to the ones occurring in winter (like class 1).

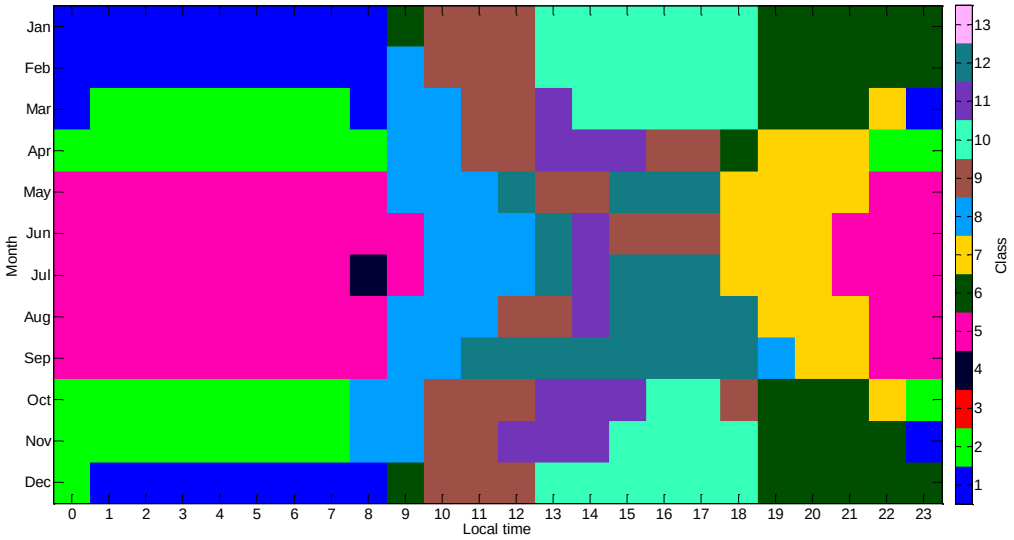

Figure 8. Most frequent class for every month and time.

To facilitate the discussion in the next section, some properties of the different classes are outlined in Table 2. The gradient Richardson number Ri is calculated using the logarithmic approximation:[15]

$$Ri = \frac{g}{\bar{T}} \frac{\Delta\theta\sqrt{z_2 z_1}}{\Delta u^2} \ln\left(\frac{z_2}{z_1}\right)$$

where the constant g(=9.81) is the gravitational acceleration and $\bar{T}$ is a reference temperature often assumed to be equal 300 K. The heights z_2 and z_1 are at approximately 100 and 75m, respectively. Ri in combination with the planetary boundary layer (PBL) height can provide an insight on the atmospheric conditions represented by the pattern.

Table 2. Mesoscale characteristic of each pattern at the center of the domain.

Class	Wind direction (°)	Ri	PBL height (m)
1	4	0.19	177
2	6	0.26	138
3	5	0.18	218
4	2	0.16	331
5	356	0.23	518
6	182	0.04	370
7	131	0.14	692
8	0	-0.34	565
9	198	-1.68	978
10	215	-0.59	1123
11	209	-1.07	1398
12	300	-0.53	876
13	343	-0.01	890

4.2. Validation of the downscaled of mesoscale fields

In this section the patterns obtained in Section 4.1 are downscaled using the microscale steady-state CFD model presented in Section 3.1. In this section the downscaled simulation results will be referred as *Coupled_SOM*. These results are validated and compared with standalone CFD simulations, which for simplicity will be referred to as *Standalone*. Comparisons between these two sets of simulations allow us to assess if the proposed framework of this study provides an improvement with respect to the typical operation of CFD models. To make this comparison possible, for each of the classes obtained with the SOM2L a directional sector is assigned based on the wind direction of the mesoscale field (Table 3). As validation metric, the absolute average of the crosscheck prediction error per class, $AXPE_k$ is used, which is defined in Section 3.4. The Standalone simulations uses 12 directional sectors using the same settings as in the study of Duran et al. (2020),[2] where the same site is validated. In Table 4, the Monin-Obukhov length and PBL height used for each directional sector is presented. It must be noted that the parameters were chosen according to the measured profiles and therefore their values might not correspond to reality. However, this reproduces the kind of decision procedure taken when doing wind resource analysis with limited information.

Table 3. Comparable results of the classes simulated by the Coupled_SOM model and the directional sectors simulated by the Standalone model.

	Comparable simulations												
Coupled_SOM	1	2	3	4	5	6	7	8	9	10	11	12	13
Standalone	0°	0°	0°	0°	0°	180°	120°	0°	210°	210°	210°	300°	330°

Table 4. Main parameters for the analytical boundary conditions of the standalone CFD simulations.

	Directional sector											
Variable	0°	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
Monin-Obukhov length	-1000	-50	-50	-50	-50	-50	-20	-20	-20	-20	-20	-50
PBL height	500	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000

The validation study is divided in the validation of the horizontal extrapolation and the validation of the vertical extrapolation. To adequately measure the differences of the models, the weighted improvement is presented, which weights the percentual error difference of the models by the frequency of the patterns and wind speed. For example, if the site has a mean wind speed of 8m/s and there is a 5% improvement for a pattern with 10% frequency and 10m/s mean wind speed, then the weighted improvement is 5% × 10% × 10m/s ÷ 8m/s = 0.625%. If

the mean wind speed of the pattern is 5m/s instead, then the weighted improvement would only be of 0.3125%. This approach allows to assess the individual contribution of the patterns considering their respective importance for the complete wind resource. The same weighting is applied when calculating the total error of the model.

Validation of the horizontal extrapolation

For the validation of the horizontal extrapolation, the measurements M1 @60m and M2 @65m are used. The validation results are presented in Table 5. The Coupled_SOM simulations show a considerable improvement respect to the Standalone simulations in terms of horizontal extrapolation. Most of the improvement comes from classes 2, 3, 6 and 8. In the case of classes 2 and 3, the error of the Coupled_SOM model is much lower; however, those classes have a relatively low wind speed, making its weighted improvement in the same order as the classes 6 and 8. The Standalone model does better in classes 4, 5 and 7, but the weighted improvement is rather modest. For classes 9, 10, 11, 12 and 13 the differences tend to be small and their impact in the overall error is negligible. In terms of atmospheric stability, the Coupled_SOM simulations tend to perform better in neutral ($|Ri| < 0.04$) and unstable conditions ($-0.4 < Ri < 0$). For very unstable conditions ($Ri < -0.4$) differences are not significant. Under slightly stable to stable conditions ($0.15 < Ri$) the relative performance of the models depends on the class. In practice, modelers would combine the best results per class. By doing this, a total error of 5.93% is achieved.

Table 5. Absolute average of the crosscheck prediction error per class $AXPE_k$ of the horizontal extrapolation for the different models.

		$AXPE_k$(%)		Weighted
Class	Frequency	Coupled_SOM	Standalone	improvement (%)
1	9.01	11.6	14.5	0.14
2	6.27	32.7	50.6	0.56
3	3.69	19.3	61.6	1.17
4	4.01	8.1	3.0	-0.21
5	7.66	12.5	11.0	-0.19
6	18.23	5.0	17.3	1.57
7	4.80	24.2	16.8	-0.31
8	20.28	0.9	3.6	0.66
9	14.11	2.9	3.0	0.00
10	4.04	2.8	1.7	-0.06
11	2.00	6.1	3.7	-0.06
12	4.43	1.9	0.3	-0.09
13	1.46	4.3	5.4	0.04
Total		6.84	10.08	3.24

As seen in Figure 9, sector 0° of the Standalone simulations (which is comparable to classes 1, 2, 3, 4, 5 and 8) has a relatively homogeneous wind field in the area of the measurements. The simulation of this sector performs similar to the class 1 of the Coupled_SOM as it also reproduces a relatively homogeneous wind field (Figure 10). However, for the classes 2 and 3, the Coupled_SOM shows a more heterogenous field, outperforming the Standalone simulations in those classes (See for example class 3 in Figure 10). Even subtle homogeneity like in class 8 of the Coupled_SOM simulation can have a big impact in the errors. However, having heterogeneity does not guarantee better results, as seen in the results of class 4, since it could be caused by noise introduced by the mesoscale model or the downscaling procedure.

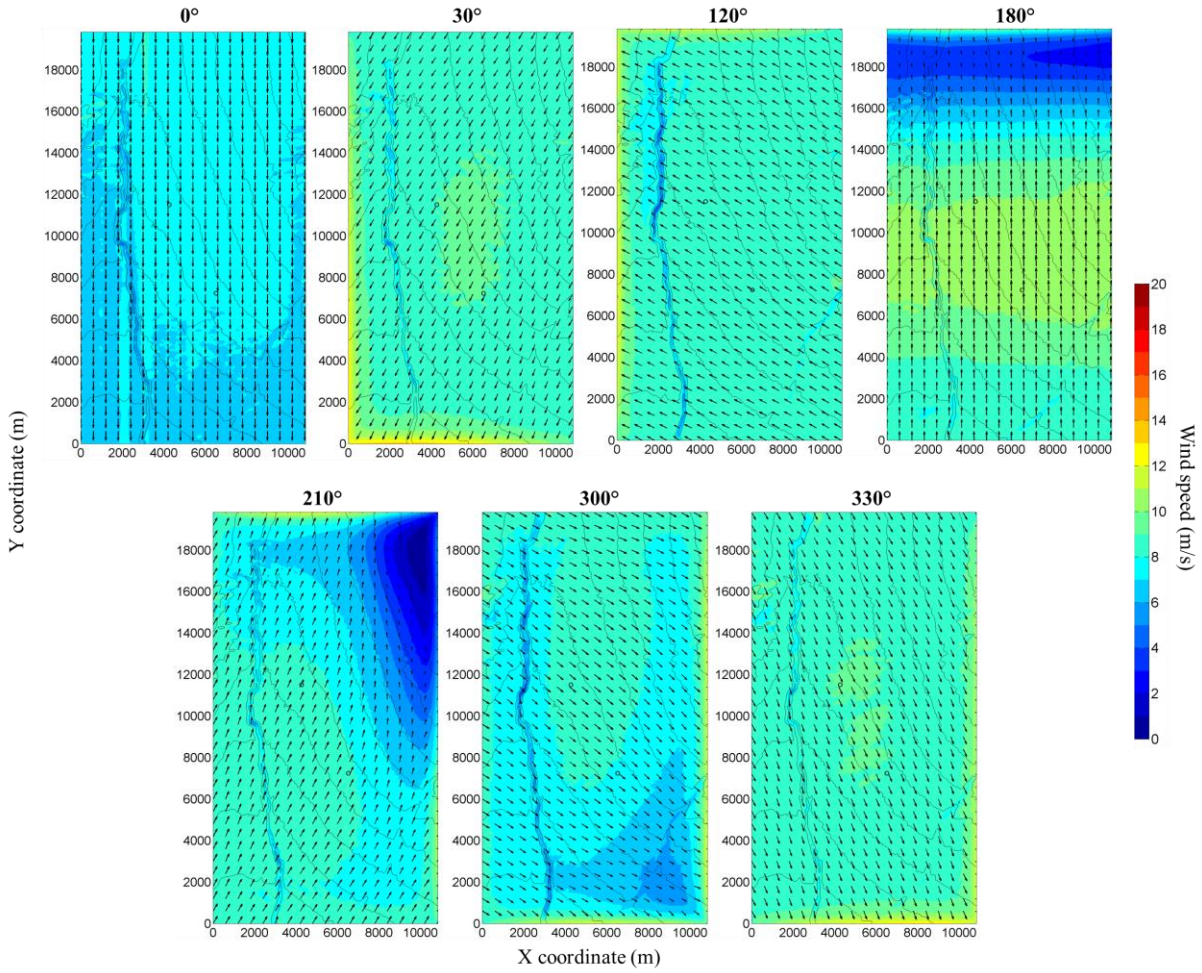

Figure 9. Horizontal planes of horizontal wind speed for some directional sectors of the Standalone simulations. The locations of M1 and M2 are indicated by a black circle.

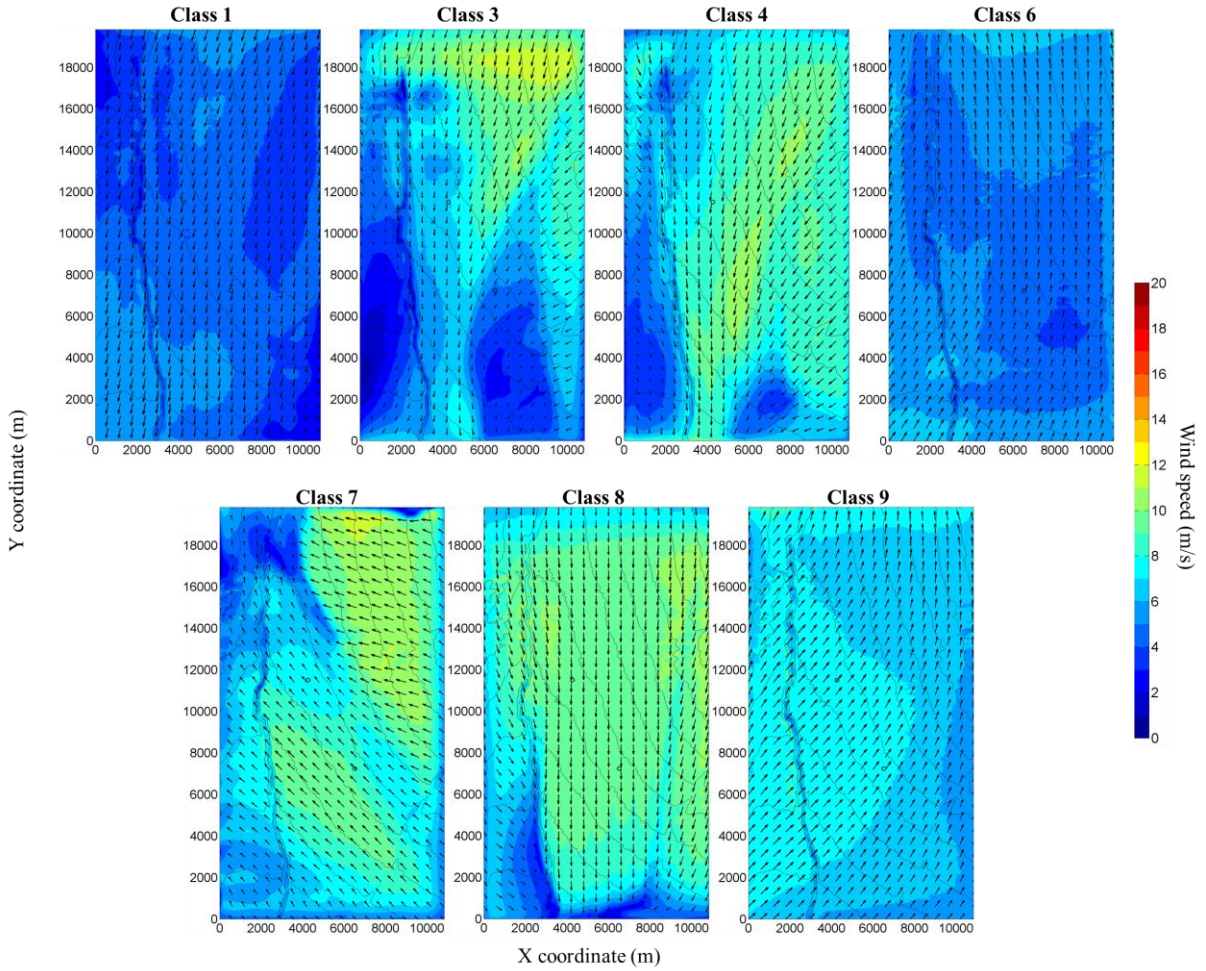

Figure 10. Horizontal planes of horizontal wind speed for some classes of the Coupled_SOM simulations. The locations of M1 and M2 are indicated by a black circle.

Similar to sector 0°, the relative performance of the models depends on their homogeneity. Most of the gaining in using Coupled_SOM comes from the more heterogenous fields as in class 6 (comparable to sectors 180° and 210°). This also means that the two models perform similarly when both have homogeneous fields as in class 9 (comparable to sector 210°). However, once again heterogeneity does not ensure better results, as seen in the case of class 7.

Validation of the vertical extrapolation

For the validation of the vertical extrapolation, the measurements of L1 @78m, 98m and 128m are used. The LIDAR measurements were chosen because it covers heights that are relevant for the development of wind energy projects. The validation results are presented in Table 6. In general terms, the difference in performance between the Coupled_SOM and the Standalone

simulations is minimal. However, it must be noted that in the study of Duran et al. (2019),[2] the parameters of the boundary conditions shown in Table 4 were chosen to fit the measured vertical profiles at M2. Given the good results of the Standalone simulations on the vertical extrapolation but not necessarily in the horizontal, it appears that the Standalone model is overfitted for the vertical profiles. Still the Coupled_SOM simulation performs similar, with slightly better results for class 8. By combining the best results, a total error of 1.79% is achieved.

Table 6. Absolute average of the crosscheck prediction error per class $AXPE_k$ of the vertical extrapolation for the different models.

		Error		Weighted
Class	Frequency	Coupled_SOM	Standalone	improvement (%)
1	9.01	8.4	6.0	-0.120
2	6.27	8.9	5.4	-0.149
3	3.69	4.6	2.8	-0.066
4	4.01	1.0	1.0	-0.003
5	7.66	1.9	1.7	-0.020
6	18.23	2.7	1.5	-0.153
7	4.80	0.9	3.3	0.099
8	20.28	2.4	3.6	0.296
9	14.11	0.8	0.4	-0.063
10	4.04	0.2	0.2	0.001
11	2.00	0.5	0.3	-0.006
12	4.43	0.6	1.2	0.030
13	1.46	1.1	0.4	-0.025
Total		2.39	2.21	-0.18

To better understand the results, the vertical profiles of wind speed at L1 are presented in Figure 11. An interesting result is that for almost all classes, the Coupled_SOM simulations reproduce negative shear at heights higher than 100m. It is clear that for example in classes 7 and 8, this negative shear better represents the measured profile, which in these cases it is reflected as an improvement in terms of error (Table 6). As seen in Figure 11, none of the Standalone simulations can reproduce this negative shear. By construction, the analytical boundary conditions used in the Standalone simulation have a logarithmic shape. Therefore, the only way for the Standalone model to reproduce a negative shear is as a consequence of the terrain, like for example at the top of a ridge. However, as shown in Figure 3 the terrain is relatively simple and therefore negative shear can only be reproduced as a consequence of weather conditions, which is achieved by using the Coupled_SOM model.

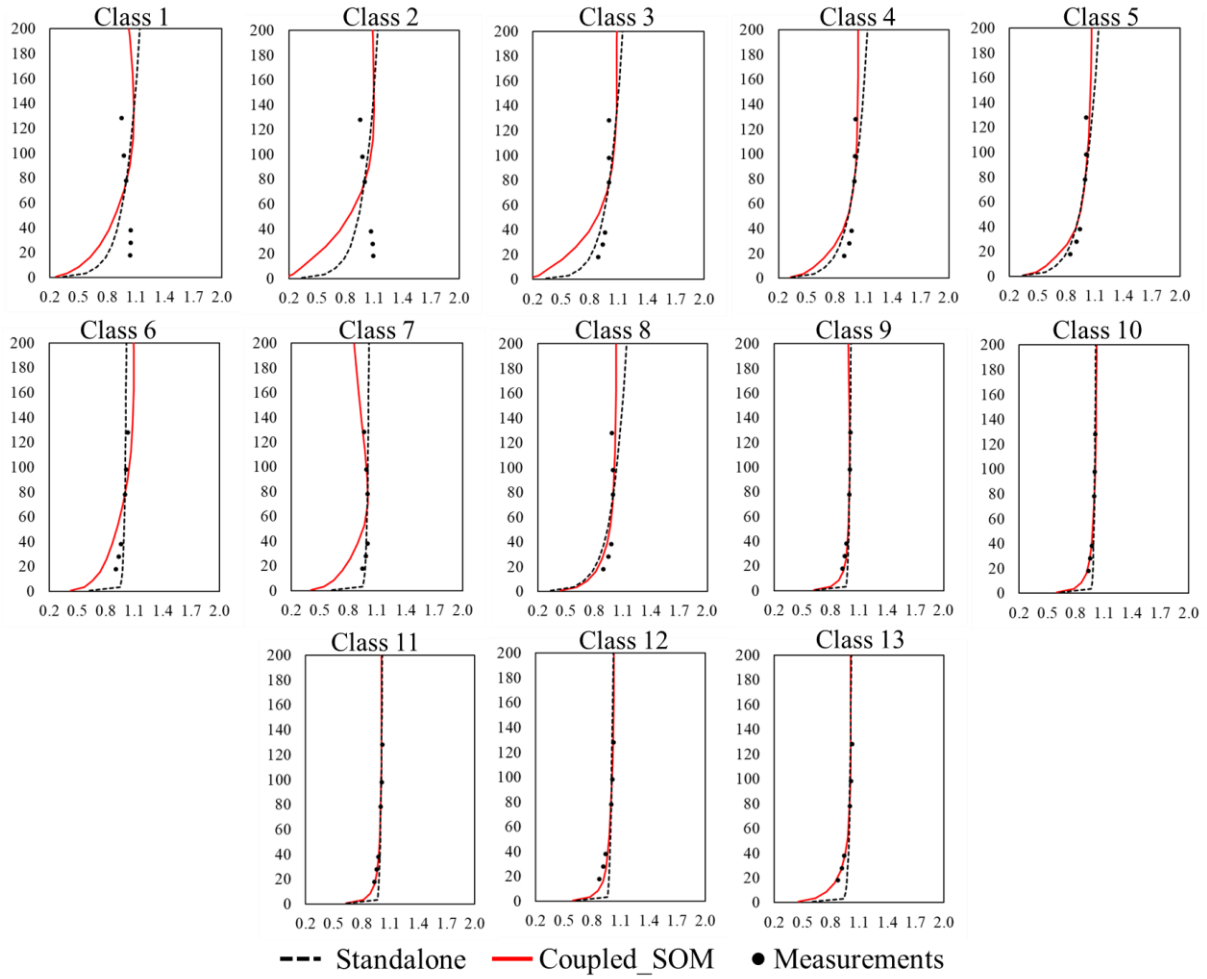

Figure 11. Measured and simulated vertical profiles of normalized wind speed at the location of L1. The horizontal and vertical axes correspond to the normalized wind speed @78m and height a.g.l. (m), respectively.

For classes 9, 10, 11, 12 and 13 the measured profiles have a shear coefficient ~0 for heights above 80m. This extremely low shear is reproduced by the Standalone model using Monin-Obukhov lengths of -50m and -20m. As commented, these values were chosen to fit the measured profile at M2. However, since the measured profiles do not follow a logarithmic shape, the fitting is conducted only for the measurements at higher heights, and therefore a mismatch at lower heights is visible (Figure 11). As seen in the same figure, this problem is not present in the Coupled_SOM simulation since the profiles do not necessarily follow Monin-

Obukhov similarity. This makes the Coupled_SOM model to be more flexible and capable to reproduce more realistic vertical profiles. However, it seems that the Coupled_SOM approach struggles to completely represent the vertical profiles for most classes with stably stratified flow like classes 1, 2, 3, 4, 5, 6 and 7.

5. Conclusions

The SOM2L classification methodology utilized in this study is able to extract the dominant mesoscale patterns in an automated and objective manner. The classification methodology is able to extract from a 3-D mesoscale dataset a variety of wind conditions without requiring any a priori knowledge from the user about the type nor the number of wind conditions at the site. Furthermore, with little effort the classification methodology provides insight on how the extracted patterns are related, how they evolve in time and how they are associated with different wind conditions like wind speed, wind direction, wind shear, wind veer or atmospheric stability. These characteristics makes the SOM2L methodology not only applicable for wind resource assessment but to be potentially beneficial for wind power forecasting. The obtained wind patterns or classes can be used as input in the predictive models as a way to account for the different wind conditions.

The meso-microscale coupling model used in this study, generally improves the wind estimations compared with the use of a standalone model at desert sites like the one studied in this work. Most of the improvement is in the horizontal extrapolation of the wind resource, mainly for stably stratified classes. In the case of the vertical extrapolation, both coupled and standalone models performed similar. However, the profiles reproduced by the coupled model seem more realistic for low or negative shear. Still there is space for improvement, especially for stable atmospheric conditions. As commented by Bleeg et al. (2015),[16] the value of some constants used in the RANS model might vary according to local atmospheric stability conditions. Further efforts to improve the proposed meso-microscale coupling methodology should focus on this path. On the other hand, further testing of the coupling approach must be conducted at other sites with different atmospheric and terrain conditions.

Assuming that a robust way to downscale the mesoscale patterns obtained by the SOM2L is available, the proposed meso-microscale coupling framework is very promising for wind resource assessment applications. One example is sites where the wind has few predominant directions. Since the predominant directions at these sites are likely to have more than one predominant wind condition, the common approach of using one simulation per directional sector is likely to be ill-suited. Another example is the CFD simulation of wake effects in wind farms. In order to resolve the wake in a steady-state CFD model (with an actuator disk for example), different wind speeds for the same wind direction must be simulated. This multiplies the number required of simulations, and by even more if different atmospheric stability classes must be taken into account. By using the SOM2L, infrequent combinations of wind speed and stability can be skipped, as the conditions derived from the mesoscale data are used instead.

Acknowledgments

This research was supported by a grant from The Norwegian Research Council, project number 271080. The author thanks to Romain Molins from Mainstream Renewable Power for granting access to onsite measurements and mesoscale datasets.

References

1. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004, doi:10.1029/2005RG000183.
2. Durán, P.; Meiβner, C.; Casso, P. A new meso-microscale coupled modelling framework for wind resource assessment: A validation study. *Renew. Energy* **2020**, *160*, 538–554, doi:10.1016/j.renene.2020.06.074.
3. Launder, B.E.; Spalding, D.B. The numerical computation of turbulent flows. *Comput. Methods Appl. Mech. Eng.* **1974**, *3*, 269–289, doi:10.1016/0045-7825(74)90029-2.
4. Gravdahl, A.R. Meso Scale Modeling with a Reynolds Averaged Navier - Stokes Solver Assessment of wind resources along the Norwegian coast. In Proceedings of the 31st IEA Experts Meeting State of the Art on Wind Resource Estimation; Risø, Denmark, 1998; pp. 1–14.
5. Meissner, C.; Gravdahl, A.R.; Steensen, B. Including Thermal Effects in CFD Wind

Flow Simulations. *J. Environ. Sci. Int.* **2009**, *18*, 833–839, doi:10.5322/JES.2009.18.8.833.

6. Vesanto, J.; Alhoniemi, E. Clustering of the self-organizing map. *IEEE Trans. Neural Networks* **2000**, *11*, 586–600, doi:10.1109/72.846731.
7. Kohonen, T.K. *Self-Organizing Maps*; 3rd ed.; Springer-Verlag: New York, 2001; ISBN 3540679219.
8. Vesanto, J.; Sulkava, M. Distance Matrix Based Clustering of the Self-Organizing Map. In Proceedings of the International Conference on Artificial Neural Networks 2002; Springer, Berlin, Heidelberg: Madrid, 2002; pp. 951–956.
9. Vesanto, J.; Himberg, J.; Alhoniemi, E.; Parhankangas, J. SOM Toolbox for Matlab 5. *Tech. Rep. A57* **2000**, *2*, 59, doi:http://www.cis.hut.fi/somtoolbox/package/papers/techrep.pdf.
10. Kohonen, T. Essentials of the self-organizing map. *Neural Networks* **2013**, *37*, 52–65, doi:10.1016/j.neunet.2012.09.018.
11. Garcia, D. Robust smoothing of gridded data in one and higher dimensions with missing values. *Comput. Stat. Data Anal.* **2010**, *54*, 1167–1178, doi:10.1016/j.csda.2009.09.020.
12. Everitt, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*; 5th ed.; 2011; ISBN 978-0-470-74991-3.
13. Han, J.; Arya, S.P.; Shen, S.; Lin, Y.-L. *An Estimation and Energy Atmospheric Theory of Turbulent Dissipation Boundary Kinetic Energy Rate Based on Layer Similarity*. *NASA technical report*; Hampton, Virginia, 2000;
14. Durán, P.; Meißner, C.; Rutledge, K.; Fonseca, R.; Martin-Torres, J.; Adaramola, M.S. Meso-microscale coupling for wind resource assessment using averaged atmospheric stability conditions. *Meteorol. Zeitschrift* **2019**, *28*, 273–291, doi:10.1127/metz/2019/0937.
15. Arya, S.P. Finite-Difference Errors in Estimation of Gradients in the Atmospheric Surface Layer. *J. Appl. Meteorol.* 1991, *30*, 251–253.
16. Bleeg, J.; Digraskar, D.; Woodcock, J.; Corbett, J.-F. Modeling stable thermal stratification and its impact on wind flow over topography. *Wind Energy* **2015**, *18*, 369–383, doi:10.1002/we.1692.