https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Identification of optimal ERA5 model level for wind resource assessments in mountainous terrain

Juan Contreras[1,2], Nicole van Lipzig[2], Esteban Samaniego[3], and Daniela Ballari[1]

[1]Instituto de Estudios de Régimen Seccional del Ecuador (IERSE), Universidad del Azuay, Cuenca, 010204, Ecuador
[2]Department of Earth and Environmental Sciences, KU Leuven, Leuven, 3001, Belgium
[3]Facultad de Ingeniería, Universidad de Cuenca, Cuenca, 010150, Ecuador

Correspondence to: Juan Contreras (juan.contreras@uazuay.edu.ec)

Abstract. Accurate estimation of hub-height wind speed is crucial for wind resources assessment at prospective sites. Traditionally, long-term wind speed series are derived from short-term site observations combined with reanalysis products, most commonly the ERA5 single-level data at 10 m and 100 m heights. However, the coarse spatial resolution of ERA5 limits their reliability in complex mountainous regions, leading to weak correlations with local wind measurements due to not adequately resolved near-surface flow. This study investigates whether the use of wind speed estimates from upper atmospheric levels (i.e., model levels) of ERA5 model level data set can improve wind speed representation in complex terrain. We compared ERA5 with hourly wind speed observations at 80 m from four meteorological masts located at high elevations (2829–3796 m a.s.l.) in the tropical Andes of southern Ecuador, and developed site-specific Random Forest (RF) models for calibrate ERA5 wind speeds. Our findings reveal that wind speeds from upper model levels (~ 1000 – 1500 m for most of the sites) exhibit substantially stronger correlations with mast observations than the theoretical hub-height. Compared with single-level inputs, model-level-driven RF estimates achieved average improvements of 59% in Perkins Skill Score (PSS), 40% in R^2, and 23% in MAE/RMSE. Importantly, the bias in annual energy production (AEP) decreased to less than 7%, in contrast with 22% when using ERA5 single-level data. These improvements were greater for sites located on exposed peaks, which are often preferred locations for wind farms, where the local flow is better captured by upper model levels. Overall, our results demonstrate that selecting appropriate upper ERA5 model levels offers a cost-effective strategy to generate accurate, site-specific hub-height wind speed time series in complex terrain. We encourage the wind energy community to exploit these upper atmospheric levels of ERA5 to enhance wind resource assessments in mountainous regions.

1. Introduction

Understanding long-term wind speed at turbine height is essential for the wind energy industry, particularly for site assessment and energy yield estimation (Watson, 2023). Direct in-situ measurements from meteorological masts remain the gold standard for assessing wind resources at onshore locations (McKenna et al., 2022; Watson, 2023). However, due to the high costs associated with deploying this infrastructure and the urgent demand for wind energy development, these measurement campaigns typically span only 1–2 years. Such short-term datasets are insufficient to capture the long-term wind conditions needed for accurate wind resource assessment over a project's lifetime (Basse et al., 2021). To overcome this limitation, the Measure–Correlate–Predict (MCP) methodology is widely used to extrapolate short-term wind speed records. MCP involves correlating measurements at the candidate site with long-term data from nearby reference sites, typically using linear regression techniques (Carta et al., 2013; Houndekindo & Ouarda, 2025b). In recent years, MCP methods have evolved to incorporate machine learning approaches for building transfer functions and have increasingly relied on wind data from Numerical Weather Prediction (NWP) models as reference sources, especially in areas lacking in-situ measurements (e.g., meteorological masts or weather stations) (Houndekindo & Ouarda, 2025b).

Data from NWP models, particularly mesoscale simulations and global reanalysis datasets—have become increasingly common as reference sources in MCP applications (Houndekindo & Ouarda, 2025b). Although mesoscale models offer higher spatial and temporal resolution than global reanalyses, they are available only for limited areas worldwide and typically cover less than 20 years (Borgers et al., 2024). This constrain is primarily due to the substantial computational resources required to run these models, limiting their utility for most parts of the world (McKenna et al., 2022). In contrast, reanalysis although with a coarser spatial

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

and temporal resolution than mesoscale models, it provides a global coverage making it useful in remote areas (i.e., non-historical monitoring sites). Moreover, they often span multiple decades (e.g., 40 to 100 years) making them suitable for long-term resources assessments. Among these, the fifth generation European Centre for Medium-Range Weather Forecasts atmospheric reanalysis system (ERA5) is the preferred reanalysis in the wind power meteorology community (Olauson, 2018). Studies by Ramon et al. (2019) and Gualtieri (2022) highlight its superior accuracy, reduced uncertainty, and greater reliability compared to other global reanalysis datasets.

Although ERA5 wind speed data is sufficiently reliable on offshore and flat areas, significant discrepancies are observed in mountainous regions. It is well-known that ERA5 tends to underestimate wind speeds—and thus wind power potential—in these complex terrains (Gualtieri, 2022). Additionally, the temporal variability of wind is reproduced less accurately compared to flat or offshore locations. These inaccuracies are largely attributed to ERA5's coarse spatial resolution, which fails to capture the intricate topography and surface roughness of mountainous areas. Grid-averaged winds in such regions often overlook speed-up effects induced by terrain features (Gualtieri, 2022). In offshore and flat locations, several studies have attempted to improve the MCP estimates incorporating physically-based covariates from ERA5 and data from mesoscale models using machine learning methods, particularly random forest (e.g., Bodini et al., 2023; Hallgren et al., 2024; Liu et al., 2023; Rouholahnejad & Gottschall, 2025; Schwegmann et al., 2023). Conversely, in mountainous areas, research in this regard is extremely limited. One notable exception is the study by Cavaiola et al. (2023), who developed a random forest model using physics-based ERA5 variables related to wind speed to estimate long-term wind power in the Alpine region. Besides, the above-mentioned studies incorporated additional meteorological or physical variables, all of them pointed out wind speed as the most critical predictor. Therefore, efforts to obtain accurate wind speed are of most relevance.

Expanding wind energy capacity, besides being strategic worldwide, is especially relevant in regions with a low diversified energy matrix. This is the case of Ecuador, which although having considerable wind resource potential in the Andes mountains, yet wind power currently accounts for only 0.6% of national electricity production (Godoy et al., 2025), revealing a largely untapped opportunity for renewable energy diversification. Additionally, the severe droughts experienced in 2023 and 2024 further exposed the vulnerability of Ecuador's electricity sector, which remains highly dependent on hydropower (Tapia et al., 2026). Expanding wind energy capacity is therefore strategic not only for diversifying the national energy matrix and increasing resilience during drought events, but also for supporting carbon-neutrality goals. In light of these challenges, there is an urgent need for robust wind resource assessment studies to guide future development. However, most potential wind farm locations in the region are situated in complex mountainous terrain, where reanalysis-based wind resource estimates remain highly uncertain.

Martinez et al. (2024) highlighted that global and regional reanalysis tend to underestimate actual site elevations due to the smoothed representation of orography in the Andes. It is known that NWP models (which form the basis of reanalysis datasets) are highly sensitive to lower boundary conditions (Hahmann et al., 2020). The degree of topographic smoothing negatively affects their ability to simulate terrain-induced wind phenomena such as anabatic and katabatic flows, mountain waves, and valley channelling which are often lost in coarse-resolution models (Kumar et al., 2025). Recent findings by Pauscher et al. (2024) indicate that reanalysis products tend to underestimate wind speeds at sites located above the grid-cell mean elevation and overestimate them at sites below it in complex terrain. This raises the question of whether wind speed estimates from reanalysis at elevations higher than the traditionally used levels (i.e., 10 m and 100 m) could provide a better match with observations in mountainous areas.

Our hypothesis is that, in mountainous regions, wind speed from upper model level in ERA5 are more representative of conditions at wind farm sites, as these locations are often exposed to free-atmospheric flow rather than surface-level wind dynamics. It is important to highlight that ERA5 single level dataset (10 m and 100 m wind components above the ground) has been stablished as a standard dataset in this field. Thus, wind speed from ERA5 single level data has been commonly used as input data for MCP studies both offshore and onshore. However, ERA5 model level dataset also provides more detailed vertical resolution data in the atmosphere (i.e., 137 model level heights) which may be more relevant for a detailed wind energy assessment. Note that this dataset has only been used in few studies, particularly at offshore locations due the higher height of wind turbines in comparison to onshore sites (e.g., Brune et al., 2021; Cheynet et al., 2025; Hahmann et al., 2022; Hallgren et al., 2024; Soares et al., 2020). Given the mismatch between

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

the orography represented in reanalysis data and the actual terrain of mountainous sites, utilizing higher-altitude wind speed data from ERA5 may be crucial for improving wind speed estimations in these areas.

Therefore, the aim of this study is to evaluate whether the use of higher atmospheric levels from ERA5 model level dataset can improve wind speed estimates and provide a more reliable time series for wind resource assessment in the mountainous regions. Our study area is the tropical Andes of southern Ecuador between 2829 and 3796 m a.s.l. To this end, we first analysed the relationship between observed wind speed time series at hub height (80 m) and ERA5 wind speed estimates across various model levels, ranging from 10 m to approximately 3200 m above ground level, in order to identify the model level that best represents wind conditions at the study sites. Next, we developed a random forest model to calibrate ERA5 and predict wind speed estimates, using the optimal model height of wind speed from the model level dataset and comparing against a reference model composed with the single level wind speed dataset. Finally, we assessed the impact of these calibrated wind speed estimates on annual energy production (AEP), applying power curves from existing wind farms in Ecuador. The findings of this study provide valuable insights for improving wind speed estimation for wind resources assessment in complex terrains using reanalysis datasets.

The structure of the paper is as follows: Section 2 describes the datasets and methodology; Section 3 presents the results; Section 4 discusses the findings; and Section 5 concludes the study with a summary and final remarks.

2. Materials and methods

2.1. Study area and in-situ masts

Our study area is located in the Andes of southern Ecuador, spanning the provinces of Cañar, Azuay, and El Oro (Fig. 1). The region is characterized by complex topography and heterogeneous land cover. High wind power potential has been identified in areas situated above 2000 meters above sea level (m a.s.l.), primarily along the ridges of the Andean Cordillera, near the Continental Divide of the Americas. This wind potential has led to the planning of several wind farm projects in the region, including the construction of Ecuador's largest operating wind farm, Minas de Huascachaca (57 MW).

The spatio-temporal variability of wind in the study area, and more broadly in the Andean region of Ecuador, is influenced by both synoptic and valley-scale circulation patterns. At the seasonal scale, wind patterns are primarily driven by the migration of the Intertropical Convergence Zone (ITCZ). The windy season, which extends from June to September, corresponds to the northward migration of the ITCZ during the boreal summer, strengthening the southern trade winds. Conversely, the calm season, from October to May, is associated with the southward shift of the ITCZ over Ecuador (López et al., 2023). Throughout the year, winds predominantly flow from the east and southeast, while westerly winds are more common during the calm season, particularly at sites located closer to the lower western flanks of the Andes. At the daily scale, thermally induced winds contribute to variability: anabatic winds occur during the day as heated air rises along mountain slopes, while katabatic winds prevail at night as cooler air descends.

Wind observations were collected from four operational meteorological masts managed by the Electric Corporation of Ecuador (CELEC EP): M1, M2, M3, and M4 (Fig. 1). M4, located in the northern part of the study area, is predominantly surrounded by cattle grazing lands and croplands. M1 and M2, situated centrally within the region, are representative of highland environments characterized by alpine grasslands locally known as *páramo*. M3, in the southwest of the study area, is also surrounded by *páramo* and pine forest. All masts are situated in wind exposed sites on ridges or hilltops.

The masts measure various meteorological variables; however, for this study, only wind speed data were used. Each mast is equipped with four first-class cup anemometers (Thies CLIMA 4.331.10.000) measuring wind speed at heights of 40, 60, 78, and 80 m above ground level, with readings recorded every 10 minutes. For our analysis, wind speed measurements at the highest level (80 m) were used. To ensure consistency between the temporal resolution of the ERA5 data and the mast observations, hourly wind speed averages were computed from the 10-minute measurements. Only complete hours, defined as those with at least six

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

valid 10-minute measurements, were included to avoid uncertainties in the hourly calculations. The wind speed observations cover the period from January 1, 2021, to December 31, 2024.

Wind speed data at 80 m underwent quality control, through detailed visual inspection and correlation analysis with wind speed measurements at the other heights, was done following MEASNET guidelines (MEASNET, 2022). Data from July 20, 2022, to September 8, 2022, at M2 were excluded due to sensor failure. Additional data gaps were identified, mainly related to maintenance campaigns and intermittent power outages caused by persistent cloudy conditions at the sites. Further details on data gaps and site characteristics are provided in Table 1.

Table 1: Details of the study sites and data availability at 80 m during the study period. The terrain category was calculated using the standard deviation of elevation within a 10 km radius around each site based on the criteria of Borowski et al. (2025).

Site	Elevation (m a.s.l.)	Land cover	Landform	SD $_{Elevation}$ (m a.s.l.)	Terrain category	Data availability (%)
M1	3796	Paramo	Ridge	471	Very complex	93.2
M2	3690	Paramo	Ridge	317	Very complex	83.4
M3	3246	Paramo/Pine Forest	Ridge	671	Very complex	99.6
M4	2829	Agriculture	Hilltop	536	Very complex	97.4

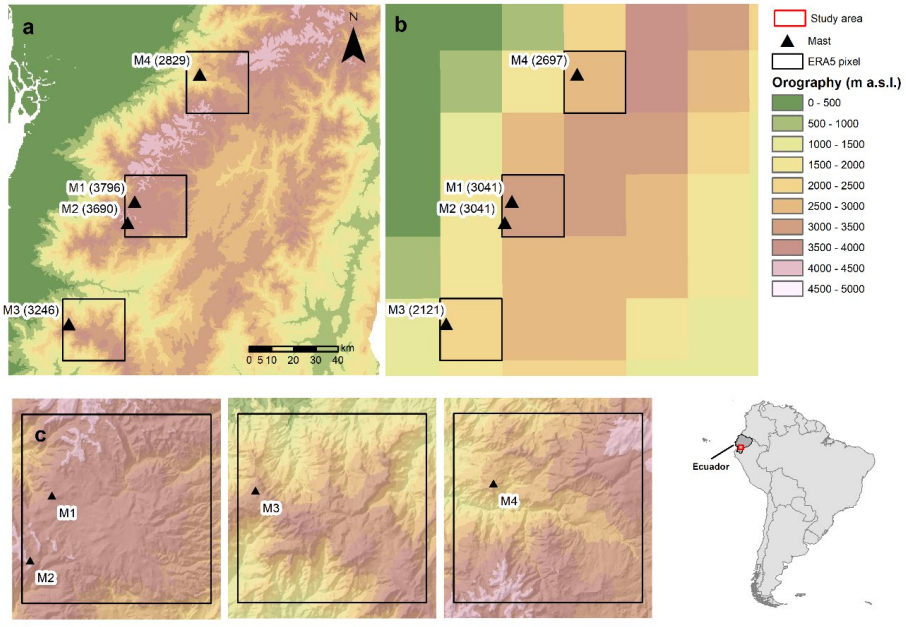

Figure 1: Map of the study area and location of meteorological masts. The orography of the study region according to the Shuttle Radar Topography Mission (SRTM) digital elevation model (30 m) and the ERA5 (~31 km) is shown in (a) and (b), respectively. In parenthesis is the in-situ and ERA5 elevation of sites in m a.s.l. The variability of orography inside the ERA5 pixel for each site is shown in (c). Grid coordinates were omitted from all maps for confidential proposes.

2.2. ERA5 data

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

ERA5 is the fifth-generation reanalysis product developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). It combines numerical weather prediction models with historical observational data using the Integrated Forecasting System (IFS) Cycle 41r2 assimilation model to provide hourly atmospheric variables dating back to 1940 (Hersbach et al., 2020). The spatial resolution of the ERA5 reanalysis dataset is approximately 31 km, with global coverage.

Two ERA5 datasets were used in this study: (1) ERA5 data on single levels, which provide wind components at 10 m and 100 m above ground level, and (2) ERA5 data on model levels, which provide wind components from the surface up to approximately 80 km altitude across 137 vertical levels. The first dataset has been widely used for wind resource assessment and is considered a standard reference in the wind energy industry, whereas ERA5 model level data have been less commonly applied in the literature.

A total of 33 model levels, ranging from 10 m (level 137 - L137) to approximately 3000 m (level 105 - L105), were selected from the ERA5 model level dataset to evaluate their relationship with wind observations. The wind components from both ERA5 datasets were downloaded using the Climate Data Store API in Python. For each of the four mast locations, ERA5 data were extracted from the nearest grid points without interpolation (see Fig. 1). This means that M1 and M2 masts have the same timeseries. Finally, wind speed was calculated from the u and v wind components for both datasets.

2.3. Evaluation and selection of optimal heights from ERA5 model level data

Measure-Correlate-Predict (MCP) methods traditionally require a high degree of correlation between wind speed observations and reference data (e.g., reanalysis datasets) to be considered suitable for wind speed prediction. This is commonly assessed using the correlation coefficient (Carta et al., 2013). We used the Pearson correlation coefficient to evaluate the relationship between wind speed measurements from the meteorological masts and wind speeds from various ERA5 model levels (ranging from 10 m to 3000 m). The model level height with the strongest correlation was then selected to estimate wind speed.

2.4. Wind speed prediction

Wind speed predictions were obtained using the Random Forest (RF) algorithm (Breiman, 2001). This method has been widely used in recent years to estimate wind speed and wind energy production due to its high flexibility and robustness. RF has demonstrated comparable results to other sophisticated machine learning models (Abdelsattar et al., 2025), and is considered one of the most popular methods for wind speed prediction (Houndekindo & Ouarda, 2025).

The RF algorithm builds an ensemble of individual decision trees. Each tree is constructed using a random subset of the training data, which minimizes overfitting and ensures independent predictions. The final prediction is computed as the average of the individual tree outputs. Additional details on the RF algorithm can be found in Breiman (2001) and Gentleman and Poggi (2020).

In this study, the RF model was preferred over a conventional regression model because it provides more consistent predictions (e.g., non-negative values), particularly at low wind speeds. Accordingly, the RF was applied to estimate wind speed at 80 m for each site, using observed wind speed at this height as reference data and ERA5 wind speed as predictors. Specifically, we trained two RF models: (1) RF1 - a model using hourly wind speed at 10 m and 100 m from the ERA5 single-level dataset, and (2) RF2 - a model using hourly wind speed at the height with the strongest correlation to mast measurements from the ERA5 model-level dataset (identified in section 2.3). The first model serves as a benchmark to compare the differences and potential improvements achieved by using the optimal model level height. It is important to note that, although most studies commonly use either 10 m or 100 m ERA5 wind speed variables to estimate near-surface wind speeds, we used both variables, as this combination explained a greater portion of variance in the RF models than using a single variable alone (see Table. A1).

For each site, both models were trained using observed hourly wind speed data from the first three years of the monitoring campaign (i.e., January 2021 to December 2023), and model performance was evaluated using data from the last year (i.e., January 2024 to December 2024). Further details on the number of samples used for training and testing for each site are provided in Table 2.

The number of trees and the minimum leaf size, which are the most important parameters of the RF model, were set to 500 and 5, respectively. These values correspond to the recommended default settings of the

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

randomForest function in R (Breiman et al., 2025). Since the main objective of this study is to highlight the improvements achieved by incorporating appropriate ERA5 wind speed heights, we did not conduct a hyperparameter optimization for each site. This decision is further supported by the fact that some studies have reported similar values for their optimal parameter settings (Hallgren et al., 2024; Liu et al., 2024).

Table 2: Number of samples used for training period (2021-2023) and testing period (2024) for the random forest models. Percentage of data available relative to each period is shown in parentheses.

Site	Training period	Testing period
M1	23890 (90.9%)	8784 (100%)
M2	20458 (77.8%)	8784 (100%)
M3	26125 (99.4%)	8784 (100%)
M4	25904 (98.6%)	8237 (93.8%)

2.5. Performance evaluation metrics

In order to quantify the discrepancies between the wind speeds estimated by the RF models and the mast wind speed data, three commonly used metrics were employed: root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). The RMSE quantifies the average magnitude of the errors and indicates how much the predictions deviate from the reference data. A higher RMSE reflects larger deviations, as this metric gives greater weight to larger errors due to the quadratic term. The MAE calculates the mean of the absolute differences between the reference values and the predictions, treating larger and smaller errors equally without applying additional weighting. The R^2 coefficient measures the strength of the linear correlation between the observed and predicted wind speeds. These metrics are defined as follows (Eqs. 1–3):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

where y_i and $\hat{y}_i$ are the ith measured and the corresponding predicted values of wind speed. The average of the measured wind speed values is denoted by $\bar{y}$. The total sample size in the test set is N.

In addition, the Perkins Score Skill test (PSS) was employed to quantifies the discrepancies in the frequency distribution between estimated and observed wind speed using Eq. (4).

$$PSS\ (H_1, H_2) = \sum_{b=1}^{n} MIN\left(F_{H_1}^b, F_{H_2}^b\right) \quad (4)$$

where H_1 and H_2 represents the first and second histogram and F^b represents the normalized frequency for bin b. The PSS represents the fraction of overlap between the two histograms, so that a PSS of 1 represents complete overlap while a value of 0 indicates a complete mismatch. The PSS was calculated considering a detailed bin width of 1 m s^{-1} similar to Borgers et al. (2024).

2.6. Wind energy estimation and evaluation

This subsection outlines the methodology used to evaluate how wind speed data from different RF models (i.e., using ERA5 single-level or model-level data) influence the estimated annual energy production (AEP) at the four study sites. Two power curves were considered in the analysis, based on the turbines installed in existing wind farms in continental Ecuador (Villonaco: 2700 m a.s.l. and Minas de Huascachaca: 1100 m a.s.l.): the Goldwind GW70/1500 and the Vestas V112/3450. Although Minas de Huascachaca operates with Dongfang Electric Corporation wind turbines, the corresponding power curves were not publicly available. Therefore, we used the Vestas V112/3450 power curve, which has similar characteristics to the turbines installed at that site. Details on the turbine operational ranges, hub heights, and the power curves used in this study are provided in Table 3 and Fig. A2 (in Appendix section), respectively.

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

The AEP was calculated using wind speed estimates from the RF models at a hub height of 80 m. The percentage error (PE) was used to quantify the differences in wind energy estimates based on the different wind speed predictions. This evaluation was carried out for the year 2024 which was not part of the RF training dataset. The PE was calculated using Eq. (5).

$$PE = \frac{AEP_{RF} - AEP_{SYN}}{AEP_{RF}} \times 100\% \quad (5)$$

where AEP_{RF} is the annual energy production modelled from wind speed estimated by the RF models, and AEP_{SYN} is the synthetic annual energy production modelled from observed wind speed data.

Table 3: Details of the wind turbines considered in the study.

Model	Power (MW)	Cut-in wind speed (m s^{-1})	Rated wind speed (m s^{-1})	Cut-off wind speed (m s^{-1})	Hub height range (m)	Rotor diameter (m)	Wind class
GW70/1500	1.5	2.5	14	25	65 - 100	70.3	IEC Ia/IIa1
V112/3450	3.45	4	12.5	25	69 - 94	112	IEC A

3. Results

3.1. Evaluation and selection of optimal heights from ERA5 model level data

The correlation between wind speed observations and ERA5 model level data at different heights above ground level (i.e., geometric altitude) is shown in Fig. 2. An inverted parabolic relationship is observed between correlation magnitude and height: correlation values increase steadily from ~0.27 – 0.74 at hub-height (79.04 m; the closest ERA5 model level height) until reaching a maximum of ~0.80 – 0.89, after which the correlation begins to decrease. The maximum correlation between observations and ERA5 data was consistently achieved at heights substantially higher (~600 - 1500 m) than the measurement height (80 m) across all mast locations. To corroborate our findings, we performed the same analysis at 3 sites with flat topography in the Coast region of Ecuador (2 located at the coastline and 1 at inner location) (see Fig. A1 in Appendix). The results showed that, at coastal and flat locations, the strongest correlations with observed wind speeds occurred near the actual hub height (< 245 m), with little differences in the correlation values between the closest level to the observations and the height of highest correlation (r <0.03). In addition, higher correlations between observations and ERA5 were observed for the inner site in comparison of coastal sites. These results using ERA5 model level data indicates that wind speeds over flat terrain is representative of hub-height conditions at coastal and inner flat sites but are not representative in mountainous areas.

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

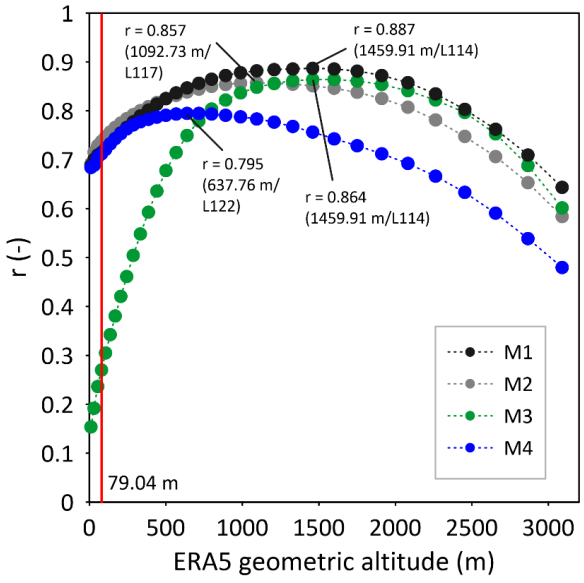

Figure 2: Correlation between observed wind speed at four sites at 80 m height and ERA5 data at different geometric altitudes for the period Jan-2021 to Dec 2024. The red line shows the closest model height to the observations. The highest correlation value and their respective altitude and model level (in parenthesis) are indicated for each site.

Notably, at M3, the correlation increased from approximately 0.27 at the measurement hub height to 0.86 at the height of maximum correlation, underscoring the potential for improving wind speed estimates by identifying optimal model heights. This pronounced difference, compared to the other sites, may be attributed to the larger discrepancy between the actual site elevation and the ERA5 model elevation. The height of maximum correlation was similar for M1, M3, and M2, while for M4 it occurred at a lower elevation. Interestingly, the magnitude of the maximum correlation was comparable for M1 (0.887), M2 (0.857), and M3 (0.864)—all of which are located on well exposed areas—but was lower for M4 (0.795), which is situated around hills probably influencing local wind speed behaviour (see Fig. 1c).

Figure 3 shows the comparison between hourly wind speed time series for the year 2024 for all ERA5 datasets and the observations. The ERA5 datasets include wind speeds at 10 m and 100 m, as well as wind speeds at the closest height to the observations (79.04 m) and at the optimal correlation heights identified in Fig. 2, both extracted from the ERA5 single level and model level dataset, respectively. Fig 3. shows substantial discrepancies between observed wind speeds and those estimated using ERA5 single level data (at 10 and 100 m). The wind speed from the model level data at the height closest to the observations shows similar variability and magnitude to the ERA5 single level data at 100 m showing a systematic underestimation around the year.

In contrast, the wind speed time series retrieved using the optimal correlation heights demonstrated substantial improvements in representing wind speed variability across all study sites, although wind speeds were still underestimated particularly during high wind periods. The dynamics of wind speed were better captured at the M1 and M2 sites compared to the others. This analysis clearly suggests the potential for improving wind speed estimates by using higher model level heights from ERA5 as reference data for MCP modelling.

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

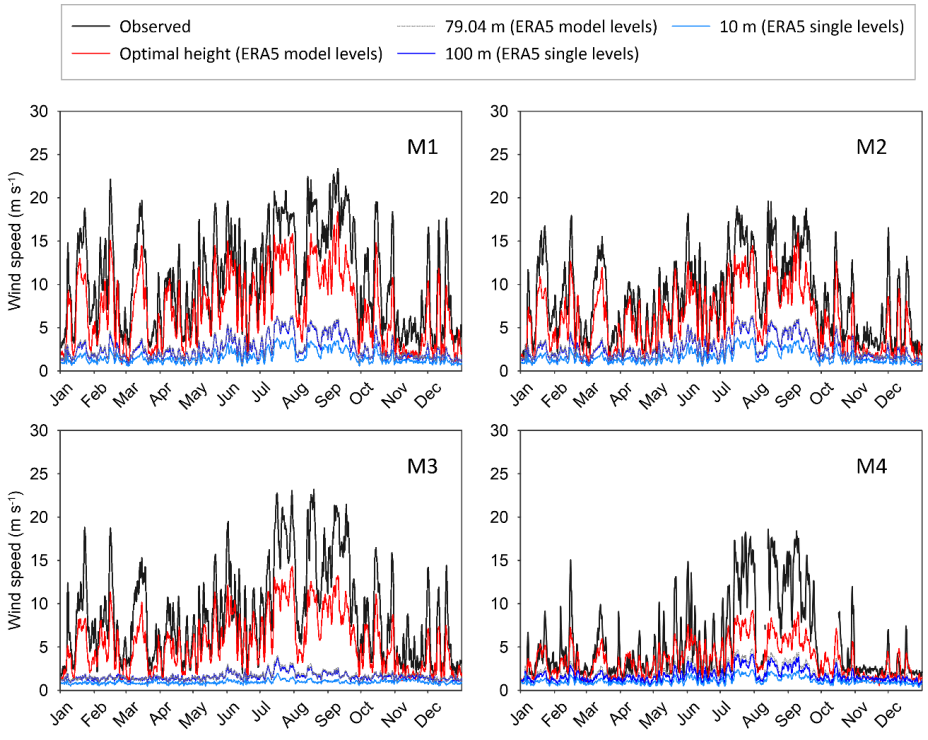

Figure 3: Wind speed time series for the observations and ERA5 data in 2024. The times series shows the 24 h running average instead of the hourly time series for visualization proposes. Note that at M1, M2 and M3, the wind speed series at 100 m (ERA5 single level) and 79.04 m (ERA5 model level) overlap.

3.2. Evaluation of wind speed predictions

The frequency distributions of observed and predicted wind speed is shown in Fig. 4 along with their evaluation metrics for the validation period (year 2024). A clear difference is observed in the measured wind speed distribution of M1, M2 and M3 compared to M4 where high frequency of low wind speed values (<2 m s^{-1}) is frequent in comparison to the other sites.

The optimal model height as input for wind speed prediction (i.e., RF2 model) provided the best match with the observed data for all sites, leading to a higher R^2 and PSS performances and lower MAE and RMSE values than the predictions that use ERA5 single level data (i.e., RF1 model). The average values of the four sites show an improvement of 40% in R^2 values while for MAE and RMSE led an improvement of 23% in both metrics. For PSS, the predictions reach an improvement of 59% respect to the reference model. These results indicate a higher improvement in the predictions to simulate the wind speed dynamics and distribution than the magnitude of wind speed.

The wind speed predictions of RF1 showed moderate similarity for all sites, with PSS values ranging from 0.75 to 0.74 at M1, M2 and M4 and 0.63 at M3. At M1, M2 and M3 it was observed an underestimation of extreme wind speed values (low and high) and the overestimation of middle range values (from 4-9 m s^{-1}). The RF2 predictions specially overcome these high discrepancies of middle range wind speed values increasing the PSS above 0.86 for all sites, indicating its superior capability to replicate the observed wind speed distributions. The improvement is especially remarkable at M3, where the PSS increases from 0.63 in RF1 to 0.90 in RF2.

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

Despite the improved performance with RF2 model, systematic discrepancies remain. In general, the modelled distributions tended to underestimate the frequency of very low wind speeds (i.e., <1 m s^{-1}) and overestimated the frequency of moderate wind speeds, particularly at M2 and M3. In addition, the frequency of high wind speeds (>15 m s^{-1}) was underestimated for all sites, especially at M1, M2 and M3. These ranges, however, varied slightly by site (Fig. 4). These biases suggested yet challenges in capturing the full variability of wind speeds, particularly at the distribution extremes.

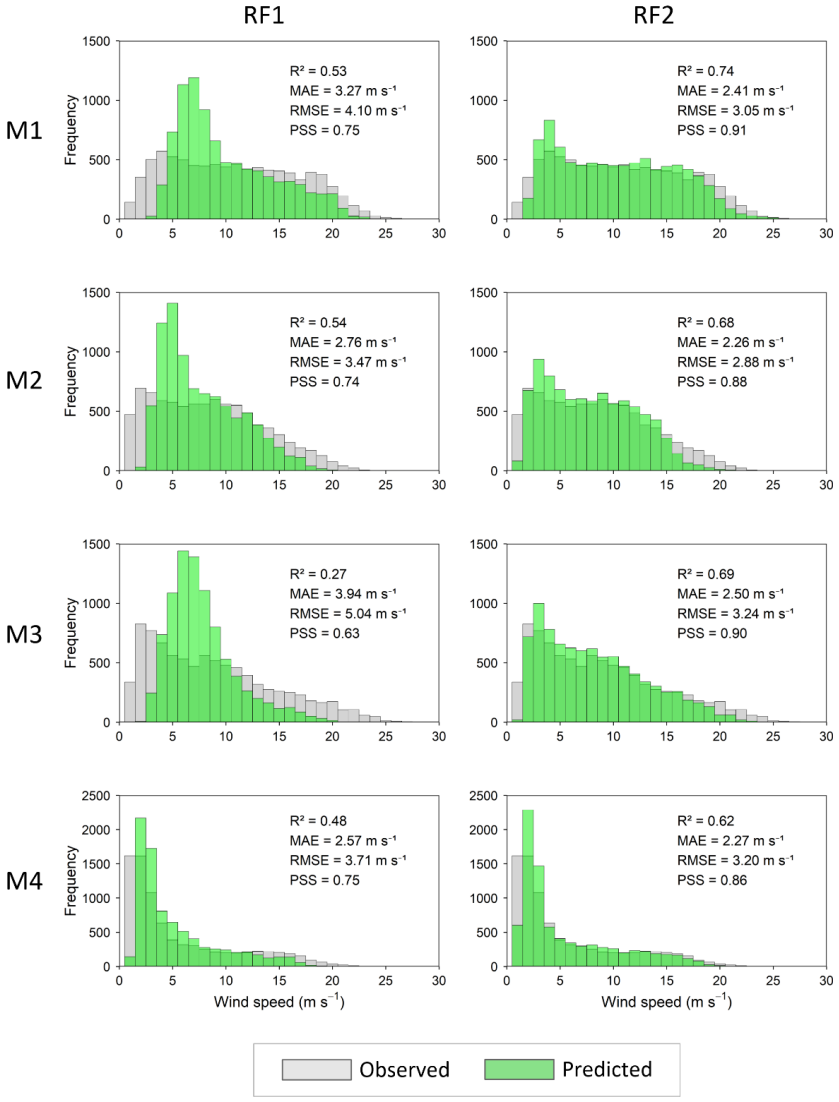

Figure 4: Comparison of wind speed frequency distributions between observed wind speeds and predictions from two Random Forest (RF) models across the four study sites for the year 2024. The left column (RF1) shows results using ERA5 wind speeds at 10 m and 100 m (single-level dataset), while the right column (RF2) shows results using ERA5 wind speed at the optimal model level height identified through correlation analysis. For each site, the figure also displays the performance metrics.

3.3. Annual wind energy estimation

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

The results presented in Table 4 show the annual energy production (AEP) estimates obtained using the Random Forest models compared to the observed AEP values for different turbine types at the four study sites. Overall, predictions using the optimal wind speed height (RF2 model) consistently outperformed the estimates based on ERA5 single level data (RF1 model), providing lower percentage errors across all locations and turbine types.

The percentage error (PE) for RF2 was reduced by approximately a factor of three compared to the reference model (RF1) for both turbines, indicating substantial improvements in AEP estimation accuracy. The similar PE values obtained using the two different power curves suggest that the choice of power curve had little influence on the relative accuracy of the AEP estimates in this case.

The RF2 model achieved the smallest discrepancies at the highest elevation site (M1), with PE values of -1.99% for the GW70/1500 turbine and -2.11% for the V112/3450 turbine. In contrast, the largest deviations were observed at M2, where the PE reached approximately 7%.

These findings highlight the importance of using appropriate ERA5 model level heights for improving wind energy production estimates in complex mountainous environments.

Table 4: Annual Energy Production (AEP) estimates from various wind speed Random Forest models compared with observed data. The best-performing models, showing the smallest discrepancies, are highlighted in bold.

Location	Turbine	AEP (MWh)	PE (%)	
			RF1	RF2
M1	GW70/1500	7787.97	-7.71	-1.99
	V112/3450	18059.60	-7.78	-2.11
M2	GW70/1500	6097.76	-16.35	-7.06
	V112/3450	14221.59	-17.60	-7.14
M3	GW70/1500	6059.96	-20.78	-5.72
	V112/3450	14063.40	-21.63	-5.77
M4	GW70/1500	3137.02	-18.83	-4.20
	V112/3450	7180.74	-20.02	-4.18

4. Discussion

Our study evaluated the hypothesis that, in the Andean mountainous region, wind speed from higher model levels in ERA5 are more representative of conditions at wind farm sites that the traditionally used 10 or 100 m single level ERA5, as these locations are often exposed to free-atmospheric flow rather than surface-level wind dynamics. Even if the topographic issue in mountain areas was previously identified in the literature (Gualtieri, 2022), this is the first time that higher model level heights were explored, identifying its suitability.

The results presented above confirms that ERA5 consistently underestimates wind speed variability in the tropical Andes in line with other studies in complex terrain (e.g., Draeger et al., 2024; Hu et al., 2023; Jourdier, 2020; Khadka et al., 2022). A central result, in line with our hypothesis, is that higher atmospheric levels of ERA5 (i.e., from model levels dataset) above the hub-height are stronger correlated to observed wind speed than lower levels at the hub-heights. This is not the case for coastal masts where higher correlations between ERA5 and observed wind speed were very close to the hub-heights (Fig. A1). Although in both Pacific coast and Andean regions, the highest correlation is above the hub-height of 80 m, these differences are significantly amplified for Andean sites. Therefore, these results support our hypothesis that observed wind speed in the Andes is more closely tied to upper atmospheric levels of ERA5 than to surface-level data.

Interestingly, optimal heights in the Andes were higher when differences between in-situ measured and ERA5 topography were larger too. This pattern may be explained by the coarse spatial terrain representation

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

of ERA5, which smooths the intricated terrain features within each grid cell, simulating lower wind speeds. For instance, the highest improvement in the level of correlation between hub height and the optimal level height was achieved particularly in M3 where differences in topography were strong. In this particular case, M3 is located at a peak compared to most of the surrounding landscape area within the ERA5 grid (Fig. 1c). These results highlight the possibility to estimate wind speed using ERA5 wind speed for a particular site by selecting higher model level heights.

A substantial improvement in the prediction of wind speed was obtained using optimal height information of ERA5 model levels in comparison with the commonly used 10 m and 100 m wind speed heights of ERA5 single levels, corroborating the suitability of using this specific dataset for mountain areas. Wind speed estimation showed similar performance for all sites using the optimal heights of ERA5 model levels compared to the ERA5 single level datasets; however, relatively small differences were noticed at M4. The lower performance in M4 is caused by the lower performance of the RF models in estimating particularly lower wind speeds (i.e., 0-2 m s^{-1}) where high frequency values within this range are common in this site.

Comparing our findings with previous studies in a strictly fair way is challenging because two main reasons: (i) most studies considered other covariables in the predictive models or/and (ii) include wind speed observations at surface level (e.g., meteorological stations) as input variables in the predictive models. It is expected that for both aforementioned cases better wind speed estimates would be achieved compared to our study that use only a single variable from reanalysis data in comparison to a model that use local information (e.g., observed wind speed generally at 2 m or 10 m). However, despite the limited input features included in our model (i.e., a single height of wind speed) and the source of its information (i.e., ERA5 reanalysis), our results were comparable in terms of performance with previous studies (Houndekindo & Ouarda, 2025a; Hu et al., 2023; Liu et al., 2023). Liu et al. (2023) used a RF model to estimate wind speed at several hub-heights using measured wind speed at 10 m (anemometer) and 300 m (radar wind profiler) and ERA5 covariables as input features for a coastal location in China. Hu et al. (2023) used the eXtreme gradient boosting (XGBoost) algorithm using topographical position index and ERA5 variables to predict wind speed at 10 m over Europe.

Our estimates were also comparable with the outcomes of more sophisticated machine learning algorithms. For instance, Houndekindo & Ouarda (2025a) used a long short-term memory (LSTM) and transformers models to bias correct ERA5 hourly wind speeds for WRA. The models outperformed static bias correction approaches and other machine learning methods. Particularly at hilly and mountainous sites, median values of the testing sites shown an $r = 0.79$. This performance is similar to that of our study, where M1 ($r = 0.86$), M2 ($r = 0.82$) and M3 ($r = 0.83$) showed even larger r values. In addition, the distribution of wind speed was better estimated with our approach showing higher values of PSS (PSS > 75% or 0.75). It is important to note that Houndekindo and Ouarda (2025) included time-resolved covariates (e.g., 10 m u- and 10 m v-components, 10 m wind speed, 2 m temperature, boundary layer height, and surface pressure) and several static covariates (derived from a Digital Elevation Model and land cover maps) in their model. All these comparisons highlight the competitiveness of our parsimonious approach for a local site estimation of wind speed in complex mountainous areas.

The study of the impact of wind estimates on energy production is not commonly assed in previous studies. The higher improvement in the reduction of underestimation of AEP estimates is promising for the evaluation of annual production. The best AEP estimates in M1 in comparison to the other sites are related to the lower occurrence of low wind speed which all the sites poorly estimate as was indicated previously. The impact of standard power curves used in this study shown negligible influence on the AEP estimation in our study area. This is because the cut-in wind speed of turbines considered in the study are above 2-4 m s^{-1}, being discarded values below this threshold in the calculation of AEP. This result indicates that wind speed estimates obtained by our approach could be reliable for the estimation of AEP using different turbine models. However, higher underestimations would be expected for sites with high frequencies of low wind speed.

It should be noted that as our main objective was to highlight the suitability of an optimal model level height to estimate wind speed, no other heights were included in the RF model. Further studies could consider additional closest level heights to the optimal level as input features to improve the wind speed estimates. We expect that the inclusion of these levels might improve the representation of interactions between atmospheric heights as they emulate wind shear effects as was evidenced using 10 m and 100 m ERA5

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

wind speed in the reference model. In addition, due to the highly complex topography interactions in the Andes, sub-grid scale variables from ERA5 representing surface-atmosphere interactions could be tested to improve wind speed estimates. Particularly, gravity waves have been identified as a relevant variable to estimate wind speed in mountainous areas (e.g., Hu et al., 2023). We are also aware that optimal heights were obtained after a detailed search of candidate heights which may require analyse relatively large amounts of information. Thus, future studies are necessary to include more masts sites to identify relationships between the performance of ERA5 and differences in topographical features as a practical model for the search of optimal levels for sites.

Anticipating the impacts of climate change on the future energy yield of wind farm projects is essential for strategic planning in the wind industry. A common approach is to directly analyse Global Climate Model (GCM) projections (e.g., Devis et al., 2018; Hahmann et al., 2022; A. Martinez & Iglesias, 2024). However, the coarse spatial resolution of GCM data (typically ≥100 km grid spacing) limits their suitability to regional or global-scale assessments. Further studies could explore the use of higher atmospheric model levels from GCMs to support downscaling of wind speed projections to local scales in mountainous regions, without the need for computationally expensive full dynamical downscaling. Alternatively, a hybrid statistical-dynamical framework, such as the approach proposed by Borgers et al. (2025), which integrates mesoscale model output with GCM projections, could offer a promising solution for producing reliable local-scale wind speed projections in complex terrain and prospective wind farm sites.

5. Conclusions

This study examined whether the use of higher atmospheric levels from the ERA5 model-level dataset can improve wind speed predictions compared to the conventional use of ERA5 single-level data in the complex mountainous terrain of the tropical Andes. Site-specific Random Forest (RF) models were trained using three years of hourly wind speed observations at 80 m from four high-altitude masts located in southern Ecuador. The predictions were validated against an independent year of observations and further tested for energy applications through the estimation of annual energy production (AEP) using two representative power curves.

The results demonstrate that wind speeds from higher ERA5 model levels (i.e., levels above ~1000 m for most sites) show stronger correlations with observed wind speeds than the conventional single-level data at 10 m and 100 m. Consequently, the ERA5 model level with the highest correlation proved more suitable for wind speed prediction. Improvements were most pronounced at well-exposed sites located on peaks, while localized sites with surrounding obstacles (e.g., M4) showed smaller gains. Predictions captured wind speed variability and distribution more effectively than absolute magnitudes (i.e., RMSE). For energy applications, the percentage error in AEP was significantly reduced to ~2–7% compared with ~8–22% when using ERA5 single level data.

These findings highlight the potential of combining higher ERA5 model level data with Random Forest models as a powerful and cost-effective approach for wind resource assessment in mountainous areas. Since this method relies on freely available reanalysis data and requires relatively low computational cost, it provides a practical alternative to mesoscale climate models for estimating long-term site-specific wind speed and energy production in complex terrain.

Appendices

Table A1: Percentage of explained variance of Random Forest models trained using various combinations of ERA5 single level wind speed inputs. In bold the best model for each site.

Input	M1	M2	M3	M4
10 m	34.02	36.01	-30.99	21.89
100 m	33.97	39.58	-16.42	31.07
10 m + 100 m	**48.04**	**50.45**	**21.78**	**48.84**

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

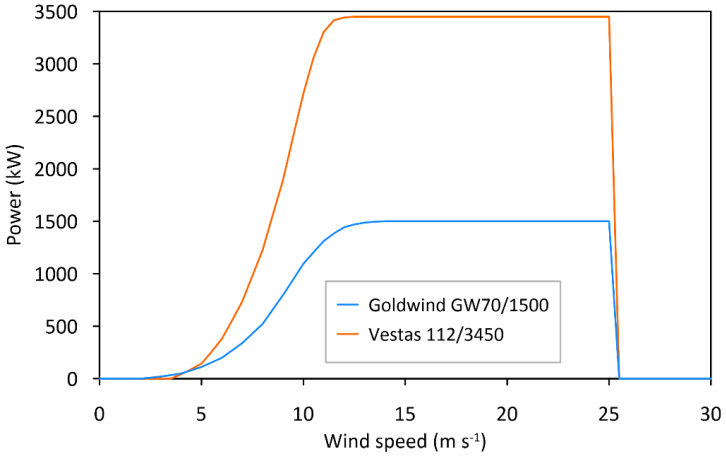

Figure A1: Power curves of the two reference turbines employed in this study.

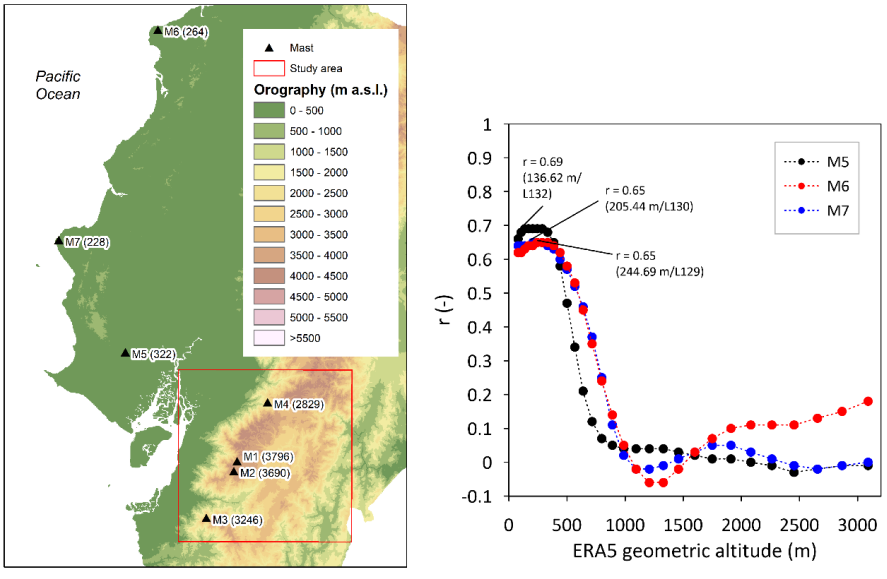

Figure A2: Correlation between observed wind speed at four Coast sites at 80 m height and ERA5 data at different geometric altitudes for the period Jan-2021 to Dec 2024. In the left side, the map shows, in parentheses, the actual elevation of each site (m a.s.l.). In the right side, the highest correlation value and their respective altitude and model level (in parenthesis) are indicated for each site.

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Code availability. The codes used for download data and analyses presented in this study are available upon request to the main author.

Data availability. Data from meteorological masts used in this research is not publicly accessible due to proprietary restrictions and confidentiality agreements.

Author contribution. JC: conceptualization, data curation, formal analysis, investigation, methodology, software, visualisation, writing (original draft preparation, review and editing); NvL: conceptualization, methodology, writing (review and editing), supervision; ES: writing (review and editing), supervision; DB: methodology, writing (review and editing), supervision, resources, funding acquisition.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors are especially grateful to CELEC for providing wind observation mast data through the inter-institutional agreement "Spatio-temporal applications supporting decision-making for renewable energy and climate change – CSR-CON-0058-23". This research was supported by Universidad del Azuay under the project “Wind energy resources in a context of change: spatio-temporal dynamics and future projections for renewable energies in areas of high topographic complexity”, grant number 2023-0178. This study is an outcome of the joint PhD agreement between the Natural Renewable Resources Doctoral Program offered by Universidad de Cuenca and Universidad del Azuay, and the Doctoral Program in Science offered by KU Leuven.

References

Abdelsattar, M., Ismeil, M. A., Menoufi, K., Moety, A. A., & Emad-Eldeen, A. (2025). Evaluating Machine Learning and Deep Learning models for predicting Wind Turbine power output from environmental factors. *PLoS ONE*, *20*(20). https://doi.org/10.1371/journal.pone.0317619

Basse, A., Callies, D., Grötzner, A., & Pauscher, L. (2021). Seasonal effects in the long-term correction of short-term wind measurements using reanalysis data. *Wind Energy Science*, *6*, 1473–1490. https://doi.org/10.5194/wes-6-1473-2021

Bodini, N., Castagneri, S., & Optis, M. (2023). Long-term uncertainty quantification in WRF-modeled offshore wind resource off the US Atlantic coast. *Wind Energy Science*, *8*(4), 607–620. https://doi.org/10.5194/wes-8-607-2023

Borgers, R., Dirksen, M., Wijnant, I. L., Stepek, A., Stoffelen, A., Akhtar, N., Neirynck, J., Van De Walle, J., Meyers, J., & Van Lipzig, N. P. M. (2024). Mesoscale modelling of North Sea wind resources with COSMO-CLM: model evaluation and impact assessment of future wind farm characteristics on cluster-scale wake losses. *Wind Energy Science*, *9*(3), 697–719. https://doi.org/10.5194/wes-9-697-2024

Borgers, R., Pinto, J. G., & Lipzig, N. Van. (2025). Quantifying the Near-Future Wind Energy Production over the North Sea Using a Novel Statistical-Dynamical Approach to GCM Downscaling. *Journal of Applied Meteorology and Climatology*, 365–381. https://doi.org/https://doi.org/10.1175/JAMC-D-24-0116.1

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Breiman, L., Cutler, A., Liaw, A., & Wiener, M. (2025). Package ‘randomForest.’ In *Machine Learning* (Vol. 45, Issue 1). https://doi.org/10.1023/A:1010933404324

Carta, J. A., Velázquez, S., & Cabrera, P. (2013). A review of measure-correlate-predict (MCP) methods used to estimate long-term wind characteristics at a target site. *Renewable and Sustainable Energy Reviews*, *27*, 362–400. https://doi.org/10.1016/j.rser.2013.07.004

Devis, A., Van Lipzig, N. P. M., & Demuzere, M. (2018). Should future wind speed changes be taken into account in wind farm development? *Environmental Research Letters*, *13*(6). https://doi.org/10.1088/1748-9326/aabff7

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

Draeger, C., Radić, V., White, R. H., & Tessema, M. A. (2024). Evaluation of reanalysis data and dynamical downscaling for surface energy balance modeling at mountain glaciers in western Canada. *Cryosphere*, *18*(1), 17–42. https://doi.org/10.5194/tc-18-17-2024

Gentleman, R., & Poggi, J.-M. (2020). *Random Forests with R*. https://doi.org/10.1007/978-3-030-56485-8_3

Godoy, J. C., Cajo, R., Mesa Estrada, L., & Hamacher, T. (2025). Multi-criteria analysis for energy planning in Ecuador: Enhancing decision-making through comprehensive evaluation. *Renewable Energy*, *241*. https://doi.org/https://doi.org/10.1016/j.renene.2024.122278

Gualtieri, G. (2022). Analysing the uncertainties of reanalysis data used for wind resource assessment: A critical review. *Renewable and Sustainable Energy Reviews*, *167*. https://doi.org/10.1016/j.rser.2022.112741

Hahmann, A. N., García-Santiago, O., & Peña, A. (2022). Current and future wind energy resources in the North Sea according to CMIP6. *Wind Energy Science*, *7*(6), 2373–2391. https://doi.org/10.5194/wes-7-2373-2022

Hahmann, A. N., Sile, T., Witha, B., Davis, N. N., Dörenkämper, M., Ezber, Y., Garciá-Bustamante, E., Fidel González-Rouco, J., Navarro, J., Olsen, B. T., & Söderberg, S. (2020). The making of the New European Wind Atlas - Part 1: Model sensitivity. *Geoscientific Model Development*, *13*(10), 5053–5078. https://doi.org/10.5194/gmd-13-5053-2020

Hallgren, C., Aird, J. A., Ivanell, S., Körnich, H., Vakkari, V., Barthelmie, R. J., Pryor, S. C., & Sahlée, E. (2024). Machine learning methods to improve spatial predictions of coastal wind speed profiles and low-level jets using single-level ERA5 data. *Wind Energy Science*, *9*(4), 821–840. https://doi.org/10.5194/wes-9-821-2024

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Houndekindo, F., & Ouarda, T. B. M. J. (2025a). LSTM and Transformer-based framework for bias correction of ERA5 hourly wind speeds. *Energy*, *328*. https://doi.org/10.1016/j.energy.2025.136498

Houndekindo, F., & Ouarda, T. B. M. J. (2025b). Machine learning and statistical approaches for wind speed estimation at partially sampled and unsampled locations; review and open questions. *Energy Conversion and Management*, *327*. https://doi.org/10.1016/j.enconman.2025.119555

Hu, W., Scholz, Y., Yeligeti, M., Bremen, L. von, & Deng, Y. (2023). Downscaling ERA5 wind speed data: a machine learning approach considering topographic influences. *Environmental Research Letters*, *18*(9). https://doi.org/10.1088/1748-9326/aceb0a

Jourdier, B. (2020). Evaluation of ERA5, MERRA-2, COSMO-REA6, NEWA and AROME to simulate wind power production over France. *Advances in Science and Research*, *17*, 63–77. https://doi.org/10.5194/asr-17-63-2020

Khadka, A., Wagnon, P., Brun, F., Shrestha, D., Lejeune, Y., & Arnaud, Y. (2022). Evaluation of ERA5-Land and HARv2 Reanalysis Data at High Elevation in the Upper Dudh Koshi Basin (Everest Region, Nepal). *Journal of Applied Meteorology and Climatology*, *61*(8), 931–954. https://doi.org/10.1175/JAMC-D-21-0091.1

Kumar, R., Rutgersson, A., Asim, M., & Routray, A. (2025). Understanding Wind Characteristics Over Different Terrains for Wind Turbine Deployment. *Meteorological Applications*, *32*(4). https://doi.org/10.1002/met.70079

https://doi.org/10.5194/wes-2025-272
Preprint. Discussion started: 22 January 2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE DISCUSSIONS

Liu, B., Ma, X., Guo, J., Li, H., Jin, S., Ma, Y., & Gong, W. (2023). Estimating hub-height wind speed based on a machine learning algorithm: Implications for wind energy assessment. *Atmospheric Chemistry and Physics*, *23*(5), 3181–3193. https://doi.org/10.5194/acp-23-3181-2023

Liu, B., Ma, X., Guo, J., Wen, R., Li, H., Jin, S., Ma, Y., Guo, X., & Gong, W. (2024). Extending the wind profile beyond the surface layer by combining physical and machine learning approaches. *Atmospheric Chemistry and Physics*, *24*(7), 4047–4063. https://doi.org/10.5194/acp-24-4047-2024

López, G., Arboleya, P., Núñez, D., Freire, A., & López, D. (2023). Wind resource assessment and influence of atmospheric stability on wind farm design using Computational Fluid Dynamics in the Andes Mountains, Ecuador. *Energy Conversion and Management*, *284*. https://doi.org/10.1016/j.enconman.2023.116972

Martinez, A., & Iglesias, G. (2024). Climate change and wind energy potential in South America. *Science of the Total Environment*, *957*. https://doi.org/10.1016/j.scitotenv.2024.177675

Martinez, J. A., Junquas, C., Bozkurt, D., Viale, M., Fita, L., Trachte, K., Campozano, L., Arias, P. A., Boisier, J. P., Condom, T., Goubanova, K., Pabón-Caicedo, J. D., Poveda, G., Solman, S. A., Sörensson, A. A., & Espinoza, J. C. (2024). Recent progress in atmospheric modeling over the Andes – part I: review of atmospheric processes. *Frontiers in Earth Science*, *12*. https://doi.org/10.3389/feart.2024.1427783

McKenna, R., Pfenninger, S., Heinrichs, H., Schmidt, J., Staffell, I., Bauer, C., Gruber, K., Hahmann, A. N., Jansen, M., Klingler, M., Landwehr, N., Larsén, X. G., Lilliestam, J., Pickering, B., Robinius, M., Tröndle, T., Turkovska, O., Wehrle, S., Weinand, J. M., & Wohland, J. (2022). High-resolution large-scale onshore wind energy assessments: A review of potential definitions, methodologies and future research needs. *Renewable Energy*, *182*, 659–684. https://doi.org/10.1016/j.renene.2021.10.027

MEASNET. (2022). *Evaluation of Site-Specific Wind Conditions*.

Olauson, J. (2018). ERA5: The new champion of wind power modelling? *Renewable Energy*, *126*, 322–331. https://doi.org/10.1016/j.renene.2018.03.056

Pauscher, L., Geiger, D., Yuan, D., Bär, F., Good, G., Spangehl, T., Kaspar, F., Weber, H., & Callies, D. (2024, September). An evaluation and comparison of wind speeds from different reanalysis models in the context of wind energy – the influence of topography. *EMS Annual Meeting 2024*. https://doi.org/10.5194/ems2024-957

Ramon, J., Lledó, L., Torralba, V., Soret, A., & Doblas-Reyes, F. J. (2019). What global reanalysis best represents near-surface winds? *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 3236–3251. https://doi.org/10.1002/qj.3616

Rouholahnejad, F., & Gottschall, J. (2025). Characterization of local wind profiles: a random forest approach for enhanced wind profile extrapolation. *Wind Energy Science*, *10*(1), 143–159. https://doi.org/10.5194/wes-10-143-2025

Schwegmann, S., Faulhaber, J., Pfaffel, S., Yu, Z., Dörenkämper, M., Kersting, K., & Gottschall, J. (2023). Enabling Virtual Met Masts for wind energy applications through machine learning-methods. *Energy and AI*, *11*. https://doi.org/10.1016/j.egyai.2022.100209

Tapia, M., Heinemann, D., & Zondervan, E. (2026). Too many eggs in one basket: On the vulnerability of the Ecuadorian power system and the need for a more sustainable and resilient strategy. *Energy Policy*, *209*. https://doi.org/10.1016/j.enpol.2025.114951

Watson, S. (2023). *Handbook of Wind Resource Assessment*. John Wiley & Sons Inc.