VORTEX

Vortex ERA5 downscaling

Validation Results

November 2017

Validation report

VORTEX FdC S.L.
Parc Tecnològic Barcelona
Carrer Marie Curie 8-14
08042 Barcelona (Spain)
Tel. +34 933 285 868

SERIES

Índex

1. Vortex system technical description ... 3
2. Validation ... 5
    2.1 Methodology ... 6
    2.2 Results ... 8
3. Conclusions ... 11

1. Vortex system technical description

The Vortex system exclusively runs WRF (Weather Research & Forecasting Model) from macro to micro-scale (up to 100m resolution). WRF is a next-generation meso-scale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs. The effort to develop WRF has been a collaborative partnership, principally by the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration, the National Centers for Environmental Prediction (NCEP) and the Forecast Systems Laboratory (FSL), the Air Force Weather Agency (AFWA), the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (FAA).

To overcome data limitations, the wind analyst community has adopted and built upon reanalysis and derived downscaled products as long-term reference alternatives. Reanalysis data are a combo of multi-sourced conventional observations (synoptic stations, balloons/sounding devices, buoys, tracks) and remote sensing covering a period from 1979 to the present day. A new generation of reanalysis products enhanced through major efforts in climate data assimilation and modeling has been released in recent years. These reanalysis databases are: MERRA-2 (NASA), CFS/CFSR (NCEP), ERA-Interim (ECMWF) and, now, ERA5 (ECMWF). An advanced assimilation technology is employed to homogenize all the input observations in space and time. Finally, a global climate model, driven by the homogenized input data, is used to produce higher resolution products containing a large range of meteorological/oceanic variables.

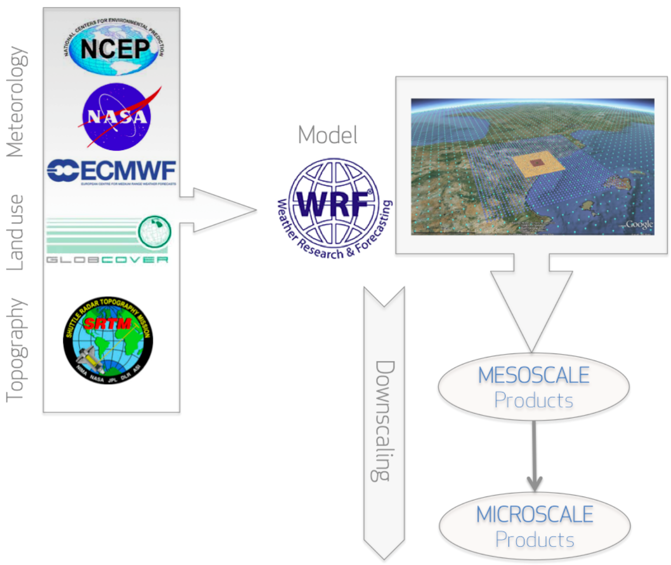

Figure 1: Schematic of the Vortex modeling process.

Vortex's implementation of the WRF model exploits a variety of global geophysical and meteorological databases (SRTM for topography data, ESA GlobCover for land cover and CFSR, MERRA-2, ERA-Interim and ERA5 for meteorological data inputs) and features a nested simulation chain ranging from hundreds of kilometers to hundreds of meters.

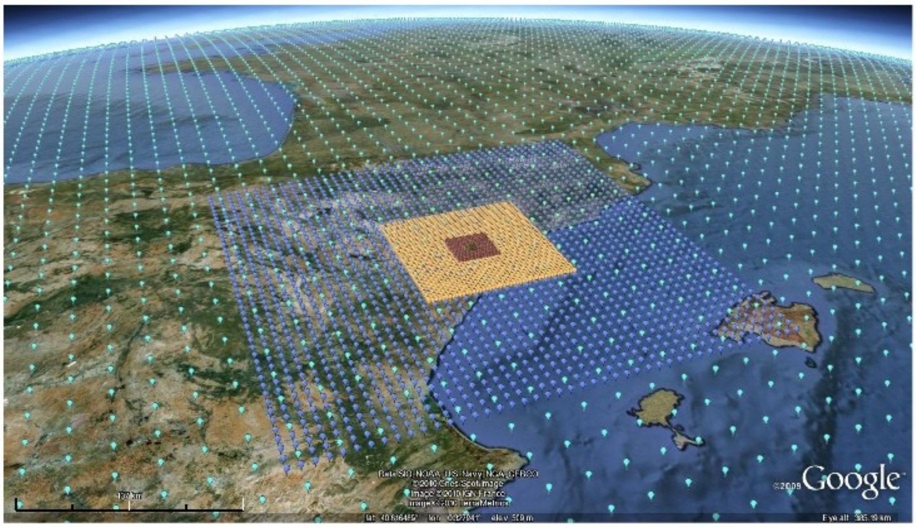

Figure 2: Downscaling illustration.

2. Validation

The present document describes a validation exercise of Vortex SERIES, where estimations of wind speed have been compared against real-world measurements.

Vortex SERIES are high-resolution time series generated using dynamic downscaling techniques driven by the large-scale conditions defined through reanalysis. Time series are computed at 3 km resolution and centered at the selected point, offering hourly values for wind speed, wind direction and other meteorological variables at different heights and for the last 10, 20 or even 30 years. By correlating the observed data to modeled time series, long-term information is added to short-term measurements to recreate historical local wind conditions.

Figure 3: Vortex SERIES output from Vortex Interface.

This validation exercise has been focused in validating Vortex SERIES generated using the new ECMWF reanalysis dataset, ERA5.

ERA5 Dataset	WRF Mesoscale Model
37 Pressure Levels + Surface	Horizontal Resolution 3km
Hourly Data	42 Vertical Levels (10 200m)
0.25x0.25 Degrees	PBL Mellor Yamada Janick
Period 2010-2016	Hourly Output
2TB per Year	* standard Vortex SERIES config

Table 1: ERA5 and WRF mesoscale model main features.

2.1 Methodology

+250 reliable measured wind-speed datasets have been used to crosscheck wind speed estimates by Vortex ERA5 SERIES. Measurements have been taken at different heights (up to 100m) and site features, located at the most wind industry interesting sites all over the world. Datasets include one or two years of 10-minute observations, in order to analyze the model performance across all seasons and have been visually quality checked to exclude any extraneous data and outliers .

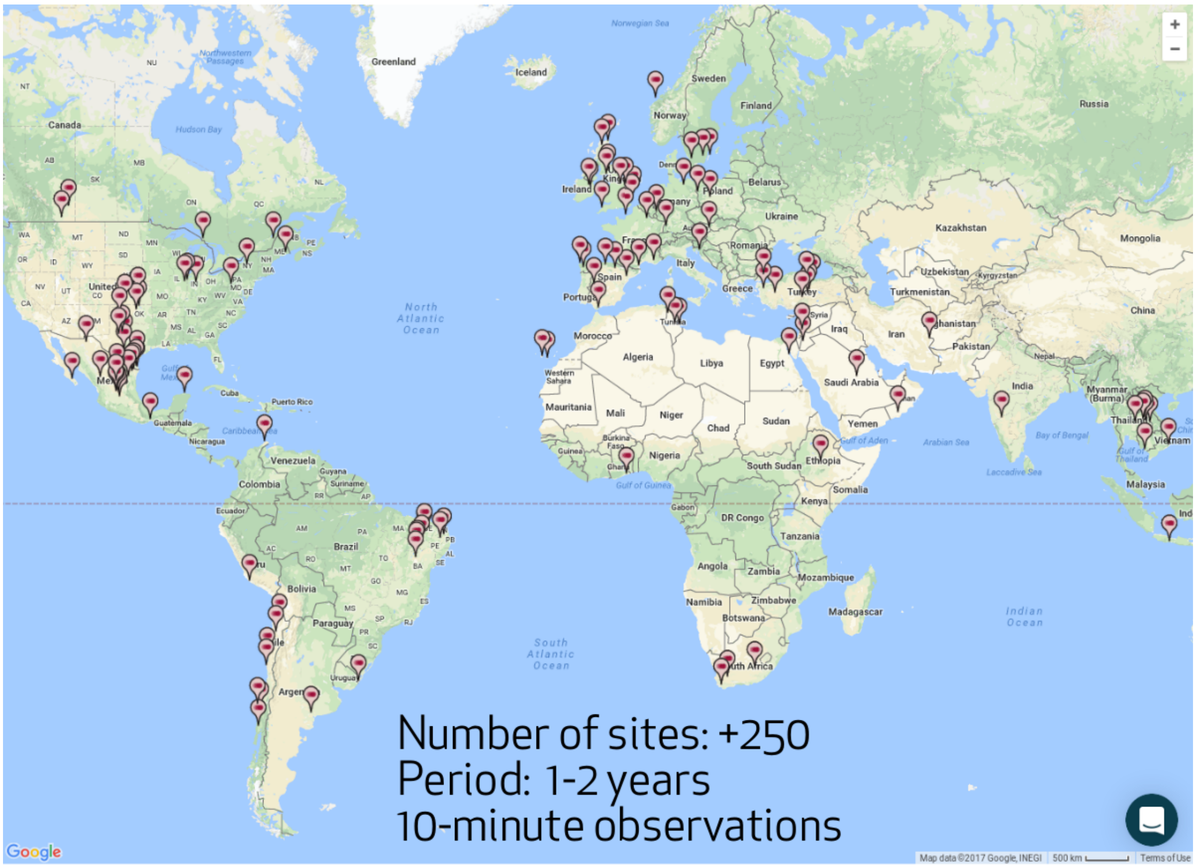

Figure 4: Measurements sites distribution .

The Following statistic metrics have been calculated to validate wind speed data delivered by WRF model runs:

- The square of the Pearson correlation coefficient, R^2, as a first raw indicator of climate representativeness. This metric represents the percentage of linear share variance between observed and modelled time series, or in other words, how much site variability can be explained (inferred) by the modelled products. High correlation values indicate better pattern similarity.

- Mean standard error or bias (in m/s and %, when normalized over mean observed wind speed) for mean wind speed, where n is the available estimations (E) and observations(O):

$$Mean\ Bias = \frac{1}{n}\sum_{i=1}^{n}(E_i - O_i)$$

- Mean absolute error for mean wind speed where n is the available estimations (E) – observations (O) . MAE is the average of the absolute errors.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|E_i - O_i|$$

- Root mean square error (m/s).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(E_i - O_i)^2}$$

- Weibull parameters error (%).
- EAP error (%).

2.2 Results

Vortex time SERIES is a product designed for long-term extrapolation of wind speed values observed in a short period of time using, for example, MCP techniques. Therefore, it is crucial to know how Vortex SERIES behaves in terms of correlation with real-world measurements. The following table and graphs show validation results when comparing measured against modeled wind speed obtained from ERA5 reanalysis sources. Wind speed correlation coefficient results are presented on the basis of hourly values and daily mean and monthly mean values.

	Vortex (default)		Vortex ERA5	
	Average	Std. Dev.	Average	Std. Dev.
R^2 (hourly)	0,59	0,14	0,67	0,13
R^2 (daily)	0,78	0,12	0,84	0,09
R^2 (monthly)	0,80	0,22	0,86	0,15

Table 2: Correlation coefficient results.

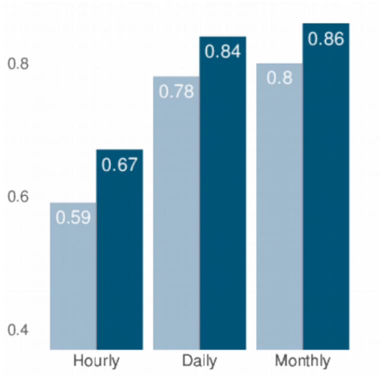

Figure 5: R² correlation coefficient: dark blue ERA5, light blue previous reanalysis.

For hourly and daily correlation, the most significant populations are shifted to greater R² in Vortex ERA5 (from 0,5-0,6 to 0,7-0,8). For hourly R², Vortex ERA5 shows population in the bin 0,9-1,00 and for daily correlation the bin 0,95-1,00 is also populated in Vortex ERA5. It can be also remark that distributions are narrower for Vortex ERA5.

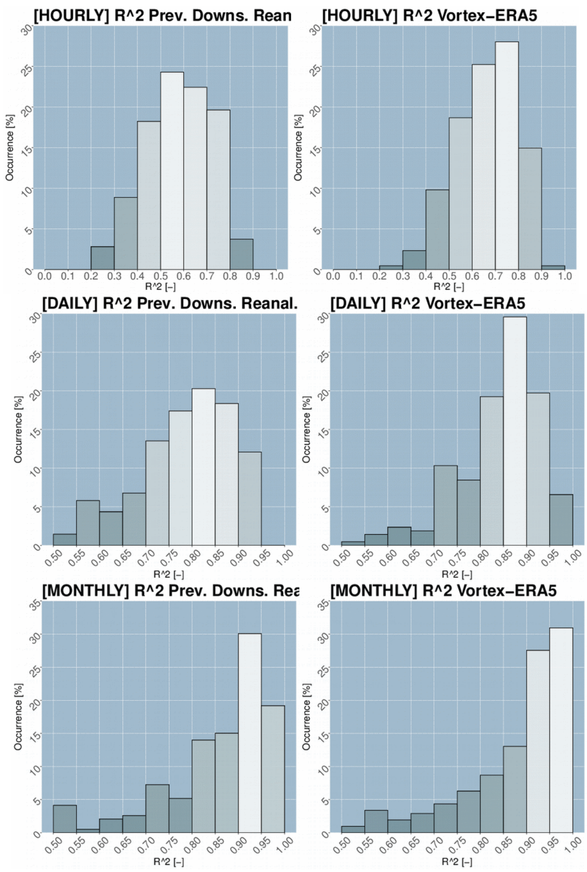

Figure 6: R² frequency of occurrence.

Following table shows the results for other statistic metrics:

	Vortex (default)		Vortex ERA5	
	Average	Std. Dev.	Average	Std. Dev.
RMSE (m/s)	2,44	0,56	2,02	0,53
MAE (%)	11,42	8,36	6,88	6,15
K-Weibull (%)	11,72	10,29	9,01	8,28
EAP (%)	26,86	20,77	16,06	12,90

Table 2: Other error results.

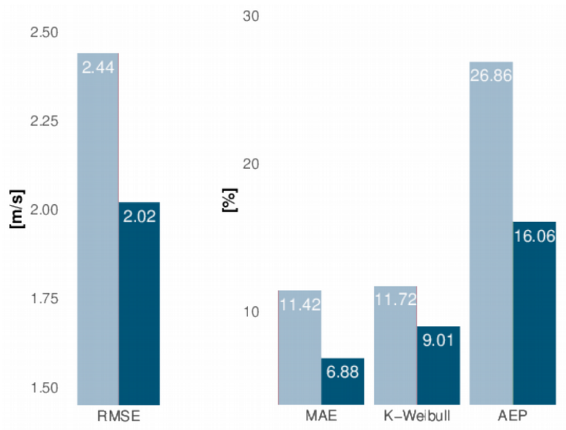

Figure 7: Other errors summary: dark blue ERA5, light blue previous reanalysis.

A very relevant improvement can be seen at all analyzed metrics when comparing Vortex SERIES ERA5 against Vortex SERIES downscaled with other reanalysis, in average.

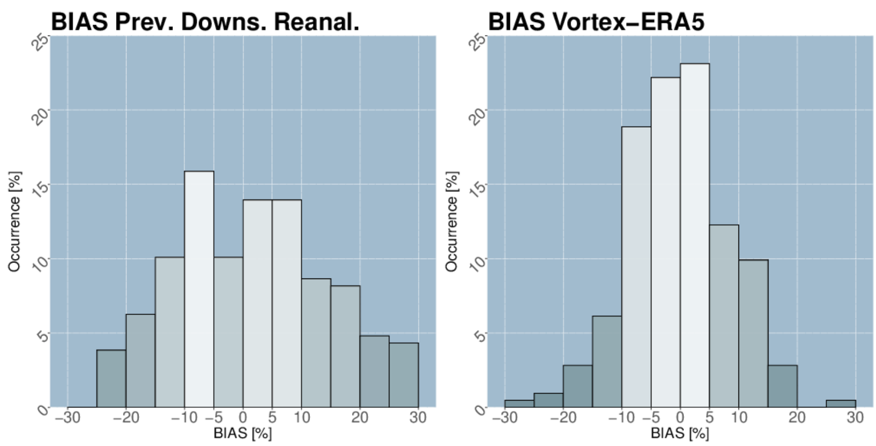

Figure 6: Bias frequency of occurrence.

The most significant populations are shifted to lower errors in Vortex ERA5 and again for Vortex ERA5 distribution becomes narrower.

3. Conclusions

The performance of long-term **Vortex SERIES** downscaled with ECMFWF latest reanalysis, **ERA5**, has been considered in this exercise. These series have been compared at **+250 locations throughout the globe**, where wind speed measurements were available at different heights and for an specific period of time (at least one full year). Modeled and measured wind speeds were compared for the same period and height. The statistics metrics used are those most common in the industry to determine the accuracy of estimates: correlation coefficient (R^2), mean bias, MAE, RSME, Weibull parameters errors and AEP deviation.

Wind-speed correlation coefficients on an hourly, daily and monthly basis have been calculated, and results show an average <u>daily correlation coefficient R^2 of 0,84</u>. Other metrics analyzed result in values of <u>6,88% for MAE</u> and around <u>2 m/s for RMSE</u>.

After analyzing +250 sites across the world, it can be concluded that **Vortex SERIES** calculated from **ERA5** reanalysis shows a **clear improvement** in terms of all statistic metrics here analyzed **versus** Vortex SERIES downscaled with **previous reanalysis**. However, <u>it will depend on the particular site which of the sources of reanalysis works best</u>.