# A risk-assessment of "wind-droughts" over India

Gangopadhyay, A. a*, Sparks N. J. b, Toumi, R. c, Seshadri A. K. d

a. Divecha Centre for Climate Change, Indian Institute of Science, CV Raman Rd, Bangalore – 560012, Karnataka, India; email. – anasuya.g.research@gmail.com ;

b. Blackett Laboratory, Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom; email. – n.sparks07@imperial.ac.uk

c. Department of Physics and Grantham Institute – Climate Change and Environment, Imperial College, Exhibition Rd, South Kensington, London SW7 2BU, United Kingdom; email. – r.toumi@imperial.ac.uk ;

d. Divecha Centre for Climate Change and Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, CV Raman Rd, Bangalore – 560012, Karnataka, India; email. – ashwins@iisc.ac.in ;

Corresponding author: Gangopadhyay, A.*

Mailing address: Divecha Centre for Climate Change, Indian Institute of Science, CV Raman Rd, Bangalore – 560012, Karnataka, India

Tel.: +91- 080-2293 3425

E-mail: anasuya.g.research@gmail.com

Formatted: Numbering: Continuous

**Abstract**

Wind power growth makes it essential to simulate weather variability and its impacts on the electricity-grid. Low-probability, high-impact, weather events , such as a wind-drought are important but difficult to identify based on limited historical datasets. A stochastic weather generator, Imperial College Weather Generator (IMAGE), is applied to identify extreme events through long-period simulations. IMAGE captures mean, spatial-correlation, and seasonality in wind speed and estimates return periods of extreme wind-events over India. Simulations show that when Rajasthan experiences wind-drought, Southern India continues to have wind and vice versa. Regional grid-scale wind-droughts could be avoided if grids are strongly interconnected across India.

**Keywords: Decarbonization,** Wind Energy, Stochastic Weather Generators, Wind Drought, Grid Interconnections

## 1: Introduction

As part of its progress toward decarbonization, India has plans to increase the share of renewable energy (wind and solar) in its electricity system[1]. Decarbonization or reduction in fossil fuel sources of energy is also important for reducing air pollution [2,3]. In an electricity grid with a large share of renewables, weather variability would impact not only demand but also the supply of electricity, making it necessary to simulate weather variability and its impacts on the power system [4]. In addition to normal weather variability, low probability high impact weather events can have an adverse impact on grid stability by creating large deficits in electricity generation. Successfully managing steep ramps in generation output, as

well as a range of demands or power generations as a result of weather variability over different time scales is crucially important for reliable operation of power systems [4].

Wind speed variability often arises from large-scale weather patterns [5], giving rise to correlated extreme excesses or deficits in wind generation. The statistics of such extreme events typically cannot be estimated from the instrumental record, because their relative short duration does not contain many realizations of such events. Nevertheless, such extreme weather events can be the result of underlying structure, such as the covariance in wind-speed between different pairs of locations. Stochastic weather generators can help in identifying such extremes, especially when they arise from statistics that can be estimated from shorter instrumental datasets.

Sparks et al. (2018) [6] presented the Imperial College Weather Generator (IMAGE), a novel multi-site multivariate stochastic weather generator that can capture various extreme events, including heatwaves and cold spells, droughts, and excess rainfall. Stochastic weather generators typically produce single-site time series of an arbitrary length of meteorological variables, while preserving statistics of the input data, which are obtained from historical observations, reanalyses, or models. In their simplest form, weather generators produce synthetic time series for a single weather variable at a single location. However, for many applications, the geographic area considered is so large that weather variables, such as wind, can vary significantly over the domain. In such situations time series at multiple sites are desirable. The production of realistic synthetic weather data, in this case, requires the preservation of spatiotemporal correlation between sites, increasing the complexity of the problem significantly, in proportion to the number of pairs of sites. Additionally, for many applications, time series of multiple, correlated weather variables are needed. The weather

generator IMAGE is designed to be used to assess the risk of events for which the spatial distribution of weather variables is essential, such as rainfall anomalies over several months over a large watershed or heatwaves affecting several regions of a country over a period of a few days [6]. This model uses multi-variate autoregressive modelling. Besides precipitation, other meteorological variables such as minimum and maximum daily temperature, solar radiation, humidity, and wind speed have been generally modelled using multi-variate autoregressive models [7]. Here we apply IMAGE for the first time to data from outside Europe and explore the risk of wind droughts across India.

## 2: IMAGE model description

We use an improved version of the IMAGE model described by Sparks et al (2018) [6]. We include only a brief description of the model here, presenting a more detailed explanation of the modifications.

All variables in IMAGE are modelled as latent Gaussian variables. At the start of simulation each variable is transformed using a normal quantile transformation such that it has a normal distribution. These transforms are performed separately for each month to allow for changes in the distribution from month to month. Once transformed, an autoregressive lag-1 model of the form

$$y_s(t) = c_s + \alpha_s y_s(t-1) + \epsilon_s \quad (1)$$

is fitted separately to each month of input data for each variable at each site, where $c_s$ is a constant, $\alpha_s$ is referred to as the memory parameter and $\epsilon_s$ is a noise term. These three parameters are each, in turn, modelled as latent Gaussian variables and are transformed such

that each parameter has a normal distribution for each variable at each site for each calendar month.

Synthetic time series are simulated for each variable at each site by first generating correlated values of $c_s$ and $\alpha_s$ for each month by sampling from a multivariate normal distribution. This process requires the decomposition of the covariance matrix of the autoregressive parameters, $\Sigma$, to a matrix $C$ such that $CC^T = \Sigma$. In Sparks et al (2018) [6] this is achieved using empirical orthogonal function decomposition, however, in this study we instead use Cholesky decomposition, which produces the same results but is computationally faster. In general, $\Sigma$ may not be positive semi-definite, which is required when sampling from the multivariate normal distribution, and therefore the nearest positive semi-definite matrix to $\Sigma$ is computed using the method of Higham (1988) [8]. Parameters are generated simultaneously for all twelve months in one simulated year, such that correlations between months in the same year are accurately simulated, as well as spatial correlation between sites. As well as simulating monthly parameters, the noise terms $\epsilon_s$ are simulated daily for each variable at each site, once again by sampling from a multivariate normal distribution. Daily values for each variable at each site can then be simulated using Eq. (1).

After simulation, variables are transformed back to their original distributions using an inverse normal quantile transformation. The pairwise Pearson's correlation coefficient of time series of variables at different sites are calculated for the simulated data and compared to the correlation coefficients of the input data. As described in Sparks et al (2018) [6], the original version of IMAGE tended to systematically under-simulate the observed spatial correlations. IMAGE has been modified to mitigate this issue using an iterative method.

Once one simulation run is complete, the covariance matrix used to generate the daily noise terms, $\epsilon_s$, is adjusted by applying a correction term equal to the difference between the observed correlation and the simulated correlation for each pair of sites. The simulation of $\epsilon_s$ is then re-run and this cycle is iterated until a satisfactorily small error in the simulated pairwise correlations is achieved. We found that ten iterations were sufficient to reach approximate convergence.

**3: Data used and Methodology**

The analysis of this paper was based on horizontal wind speed at 100 m over surface from ERA5 reanalysis products by the European Centre for Medium-Range Weather Forecasts (ERA5) [9] for 41 years (1979 - 2019) over the Indian region. We choose 100 m above surface because the hub heights of various onshore wind turbines lie roughly at that level. The spatial resolution of the reanalysis dataset is 0.25 x 0.25 degrees, and the temporal resolution is hourly. In-situ, hourly 100 m wind measurement data of 40 weather stations from National Institute of Wind Energy (NIWE) is compared with ERA5 data to check if ERA5 could capture the overall wind speed pattern.

The methodology of the paper is the following. First, we test if ERA5 could capture the overall pattern of observed wind speed over India (from NIWE measurement for 40 locations - in Figure 1, purple dots). However, we note that the goal of the paper is to present and highlight the value of the stochastic weather generator to model wind-power variability. The IMAGE model which is presented here can be used with different sources of data input, and we have chosen the best physically consistent gridded data set available to us. The input data

147 can change , but the tool presented here remains applicable to a variety of weather variables,
148 depending on the context.
149
150 Next, the IMAGE weather generator was validated for wind speeds over the Indian region.
151 The model was simulated for 40 locations (in Figure 1, purple dots) over wind-rich regions
152 located in western and southern India for daily wind speed of 4100 years. The output of 4100
153 years of IMAGE simulation for each of these locations is segregated into 100 ensemble
154 members, each of 41 years length (i.e., same length as input data). For each of these ensemble
155 members, the 41 years' daily wind speed time-series was compared with the corresponding
156 time-series in the ERA5 input dataset. The parameters chosen for comparison are the yearly
157 mean of the daily wind speed, seasonal variation of the monthly mean wind speed, and spatial
158 correlation of wind speed. Three locations are used to illustrate the validation of seasonal
159 patterns (Figure 1, green triangles), from three different high wind resource regions in India:
160 Dhanuskodi in the off-shore region near south India, and two onshore locations in Box A
161 (Rajasthan) and Box B (South India). We also examine if the wind speed simulated by
162 IMAGE could preserve the Weibull distribution of wind speeds.
163
164 As part of validation of the simulations from IMAGE, we also examine whether these capture
165 the low probability extreme events of high and low wind speed. This analysis compares the
166 return periods of various wind speed events from the ERA5 reanalysis with IMAGE
167 simulations. This validation is done for two high wind-resource regions within Rajasthan
168 (Box A) and South India (Box B) (Figure 1,brown squares).
169
170 Following the validation as described above, IMAGE has been used to simulate 1000 years of
171 wind speed over India based on the ERA5 reanalysis as input data. The ERA5 data has been

upscaled to 1 x 1 degree resolution, prior to serving as input to IMAGE. Based on the IMAGE simulations, we estimate the probability of low wind ("wind drought") over different regions and all-India. We estimate the fraction of days for which the average wind speed is below 3 m/s all over India, given one of the wind-rich regions, Box A (Rajasthan) or Box B (South India) has an average daily wind speed below 3 m/s. The 3 m/s threshold is considered because as most of the turbines have a "cut in" speed of 3 m/s, i.e., the wind speed beyond which the turbines start producing electricity.

Finally, through a case study, we illustrate the benefits of having grid-connected wind plants located in different regions, as compared to the absence of any interconnection between regional grids. These benefits are assessed from the perspective of "wind drought" or "no generation days" based on the IMAGE simulation of 1000 years. We simulate wind turbines in four sets of locations from Box A and Box B for the case study.

Set 1: One grid having maximum mean wind speed in Box A and another neighbouring grid
Set 2: One grid with maximum mean wind speed in Box B and another neighbouring grid
Set 3: Grids with maximum mean wind speed in both the boxes
Set 4: Grids with second highest mean wind speed in both the boxes

We simulate one 2.1 MW wind turbine at each location. The manufacturer's power curve of Suzlon S.88-2100 [10] model turbine with a rated power output of 2.1 MW is used for converting the wind speed to the wind generation. A lookup table created based on the normalised power curve is used for converting the IMAGE simulated wind speed to wind generation. Due to normalisation the rated generation from the turbine is indicated as 1.

Hence, wind generation for any given wind speed can be interpreted as fraction of the rated generation obtained at that wind speed.

To compare the benefits of having wind farms located in the different sets of locations, we estimate the fraction of days in 1000 years for which there was no generation from the individual plants as well as no generation from the combination of the two simulated wind-power plants. Improvement in the fraction of "no generation" days can be used as a potential parameter to assess the benefits of aggregating wind plants from different regions.

Next, we assess the robustness of this parameter (fraction of "no generation" days). Twenty grids were selected from Box A and Box B, which have higher average daily wind speed compared to other grids. Equal number of grids (10) were selected from each boxes. We simulate all possible combinations of 4 grids that can be selected from among these 20 grids. One wind turbine is simulated at each location. We estimate the reduction in percentage of "no generation" days in the aggregate generation for each combination. The reduction values are estimated based on Equation 2.

$$R = \frac{\sum_1^4 f_i}{4} - f_c \quad (2)$$

Where $f_i$ and $f_c$ are fraction of zero generation days estimated for individual and combined generations from simulated wind plants respectively. R is the absolute value of reduction in fraction of zero generation days achieved by aggregation.

## 4: Model validation

### 4.1 ERA5 reanalysis dataset and observation

Before using ERA5 as input for IMAGE, we check if ERA5 could capture the wind speed over India. Figure 2A compares the daily mean wind speed for 40 NIWE wind monitoring stations in wind rich regions of India with ERA5. There is some underestimation of point observation of wind speed in ERA5, which might result from spatial averaging of wind speed in the gridded dataset. ERA5 captures the spatial pattern of daily mean wind speed with moderate accuracy (correlation coefficient is 0.58 with p-value ≈ 0). However, it is important to note that the focus of the paper is to present and highlight the value of the stochastic weather generator to model wind-power variability , and the tool presented here can be applied to a variety of weather variables.

ERA5 outperforms ERA-Interim and MERRA-2 consistently across many parts of the world and so this reanalysis is recommended for local wind power studies[11]. Molina et al (2021)[12] show that ERA5 is able to reproduce the wind speed spectrum over Europe. The authors note[12] that despite shortcomings ERA5 provides a regular spatial and temporal wind distribution that is important for renewable energy studies. Belmonte and Stoffelen (2019)[13] analyze differences between ERA-Interim and ERA5 surface winds fields relative to Advanced Scatterometer (ASCAT) ocean vector wind observations and note that ERA5 performs better than ERA-Interim in terms of mean and transient wind errors.

Figure 2B compares observed and ERA5 monthly wind speed patterns for a NIWE wind monitoring station, Devereddypalli; The red dot indicates the monthly median wind speed

from 41 years of ERA5, the vertical line shows the distribution; Although, ERA5 underestimates the wind speed, the seasonal pattern is captured accurately.

**4.2 Mean wind speed, seasonal pattern and spatial correlation**

Simulations from IMAGE are able to represent the mean wind speed very well over India. The validation IMAGE simulations are performed for 40 locations each having 100 ensemble members that are as long as the input data, i.e., 41 years. A linear regression between the mean wind speed from the input data and IMAGE simulations was performed, and the bias is small, and the R square value is 0.99. This comparison is made in Figure 3A, where the dots show the median values and the vertical lines indicate the distribution of the ensemble members for each location. This distribution is narrow, indicating that each of the ensemble members approximate the mean wind-speed closely.

Next, we assess if IMAGE simulations could represent the spatial correlation of wind speed between various locations. We consider Dhanuskodi as a reference location, and compare the spatial correlation coefficient between daily wind speed for ERA5 input data and IMAGE simulations between Dhanuskodi and each of the other 40 locations (Figure 3B). Each of the dots in Figure 3B indicates the median of the 100 ensemble members from simulation plotted against the ERA5 value, between Dhanuskodi and one other location. The vertical lines show the distribution among the 100 ensemble members of the correlation coefficient, for each pair. These results demonstrate that IMAGE is able to successfully represent the spatial correlation of wind speed in its simulations. The R square value of a linear regression between the correlation coefficients of input and simulated output is 0.99, and the bias is small.

Much of India experiences high wind speed during the summer monsoon months (JJAS – June, July, August and September). Figure 3C shows seasonal variation of wind speed at three locations indicated in Figure 1: Dhanuskodi, Devereddypalli and Bassi. For each location, the solid lines indicate the ERA5 inputs while the dashed lines indicate distributions from the IMAGE simulations. The crosses and the triangles show the ensemble medians of the monthly wind speed from ERA5 and IMAGE simulations respectively, and the vertical bars indicate the distribution in each case. The IMAGE simulation captures this seasonal variation quite well.

**4.3 Wind speed distribution pattern**

Wind speeds are expected to follow a Weibull distribution[14]. Figure 4 compares the probability distribution plot of the ERA5 input data for 41 years and IMAGE simulated wind speed for 4100 years (4100 x 365 data points) for two locations, Gudaparihar and Bassi. The red lines show the Weibull distribution fits with appropriate shape and scale factor for the distributions. This shows that IMAGE simulations could preserve the wind speed distribution.

**4.2 Return Periods**

We estimate the return periods of different values of wind speed for both the wind rich regions (Box A and Box B) shown in Figure 1. Figure 5A and Figure 5B show the return periods of different values of spatially averaged wind speed in Box A (Rajasthan) and box B (South India). The red dots show the median of the return period based on ERA5 input data

(41 years) while the blue dots show the simulated return period in 1000 years. Corresponding ranges are also indicated by the horizontal bars. The results show that the ERA5 dataset does not contain many instances of very high wind speed, owing to its limited length, and hence could not predict the return period values for these cases. However, from the IMAGE simulation such return periods can also be estimated. A similar result occurs for the case of very low wind-speed.

This analysis of return periods is validated by the observation that the relatively frequent events with return periods much smaller than a year have similar distributions in both the ERA5 reanalysis and IMAGE simulations (Figure 5A and B). Furthermore, in the past 41 years, in the ERA5 reanalysis Box A had a record average daily wind speed of 12.25 m/s only once and the median return period of this high wind speed is predicted as 33 years by IMAGE. Similarly, Box B experienced only one record instance of average wind speed beyond 12.25 m/s in 41 years in the ERA5 dataset and the median return period of this event is predicted as 60 years by IMAGE. For Box B IMAGE can simulate extreme low wind speed of 1.25 m/s with a return period of 30 years. This is outside of the observed range and illustrates the benefits of the IMAGE model.

## 5: Results

### 5.1 Probability of low wind or "wind drought" in India

In a future electricity grid in which the share of wind energy is large, the probability of low wind ("wind drought") across large parts of the grid is a critical concern. Rajasthan and South India significantly comprise India's two distinct wind rich regions. Since wind development

316 is likely to concentrate substantially in these regions, we examine the association between
317 low wind days in both of these regions. IMAGE simulations over a period of 1000 years
318 indicate that, on days when Box A (Rajasthan) experiences low wind on average (below 3
319 m/s), there is only 0 to 10% probability that locations in Box B (South India) will have wind
320 lower than 3 m/s (Figure 6A). Low wind in Box A occurs in 7% of the days in 1000 years.
321 Similarly, on the days when Box B has low wind, with average below 3 m/s, there is only 10
322 to 20% probability that individual locations within Box A will also have low wind (Figure
323 6B). Low wind in Box B occurs 6% of the time in 1000 years. Clearly, these two regions
324 demonstrate a complementary behaviour from the perspective of "wind drought". Hence,
325 there is a possibility of avoiding grid-wide wind droughts if regional grids in these two
326 regions are themselves connected. This is illustrated further with the help of a case study.
327
328 **5.2 Case Study**
329
330 Four sets of paired locations from Box A (Rajasthan) and Box B (South India) are chosen for
331 the case study. As Box A and Box B demonstrate a complementary behaviour from the
332 perspective of "wind drought", these sets of pairs are chosen accordingly. Daily generation
333 from one wind turbine was simulated for each location depending on the local wind speed.
334
335 Set 1: One grid having maximum mean wind speed in Box A and another neighbouring grid
336 Set 2: One grid with maximum mean wind speed in Box B and another neighbouring grid
337 Set 3: Grids with maximum mean wind speed in each of the boxes
338 Set 4: Grids with second highest mean wind speed in each of the boxes
339

We estimate the fraction of days with zero generation in the 1000 year IMAGE simulations, for the individual plants and with the combination of the two plants, one in each of the chosen grids. In the case of Set 1, the individual plants in Rajasthan have 11% and 15% of days with zero generation, while the combination of these two plants has 10% of days with zero generation. Similarly, in the case of Set 2, the individual plants in South India have 11% and 9% of days with zero generation, whereas the combination of these two plants has 5% of days with zero generation (Table 1). These two case studies demonstrate that in case of wind plants situated nearby, such as in a neighbouring grid, their aggregate generations do not show much improvement as measured by the fraction of days with zero generation. This occurs because low wind days tend to coincide for the grids that are located in the same box or within small regions.

The next two sets (3 and 4) show the benefits that can be achieved by combining wind plants that are located in different regions (Box A - Rajasthan and Box B - South India). In the case of Set 3 (Combination of grids with maximum mean wind speed from each of the boxes A and B) we estimate the fraction of days with zero generation in the 1000 year IMAGE simulations for each individual locations as well as their combination. We found that while the plants in Boxes A and B experience 11% and 9% of days with zero generation, the combination of these two plants experiences only 1% of days with zero generation. Set 4 considers the combination of grids with second highest mean wind speed from each of the boxes A and B. Similar estimations were made for set 4. We found that for Set 4, while the individual plants have 15% and 11% of days with zero generation, their combination has only 2% of days with zero generation (Table 1). In the cases 3 and 4, improvement in fraction of days with wind drought for the combined generation is evident because low-wind days in each of the two regions (A and B) coincide less frequently.

The next part of the case study further examines the reduction in fraction of "no generation" days, and Figure 7 depicts the results. After identifying the 20 grids with highest mean wind-speeds in Boxes A and B, we simulate the aggregate generation from all possible combinations of 4 grids chosen from these 20. The boxplot indicates the reduction in fraction of zero generation days in 1000 years due to aggregation of different combinations of four wind plants from different Boxes (A and B); Combination AAAA has all four grids from Box A, while combination AABB has two grids from each boxes (Box A and Box B). The bold black line is the median for the improvement for each set of combinations. The top and bottom of the boxes indicate the $75^{th}$ and $25^{th}$ percentile values for each combination. Combination of plants from different regions (AAAB, AABB, BBBA) demonstrates larger reduction (0.13 to 0.14 (median)) in fraction of "no generation" days by aggregation of generation compared to plants located in same region (AAAA, BBBB) (reduction of 0.1).

**6: Discussion and Conclusion**

The IMAGE weather generator has been validated for wind-power studies over India. The weather generator has been able to reproduce the statistics of ERA5 reanalysis over the region. We tested the output for 40 grid locations, chosen because of their proximity to in-situ wind measurements from NIWE. The weather generator could correctly reproduce the mean wind and seasonality. The key advantage of the underlying model is that it can also capture the pairwise temporal correlation between sites. We confirmed that the correlation between sites is correctly captured by IMAGE. The model was then trained on wind-speed time-series from ERA5 reanalysis wind data to gain insight into the correlated behaviour of wind-droughts in the important wind resource regions of Rajasthan and South India. Our emphasis

in this paper is on noting the value of a new tool for studies of renewable drought over the region, using the relevant gridded datasets, not limited to ERA5. Other datasets, such as the regional high-resolution analysis, IMDAA, can also be used as inputs to IMAGE. The benefit of a stochastic weather generator is that it can simulate out-of-sample events to get a more robust estimate of, for example, 100 year and other low probability events. A 1000-year simulation of daily wind-speeds allows us to quantify the likelihood of wind droughts anywhere in India.

A potential implication of such studies is to that they can be used to quantify benefits of strong grid interconnections across weakly correlated regions. In our case study, we find that the risk of a wind drought in one region could be substantially mitigated by supplying wind from another region. We find cases where the number of days with no power generation can be dramatically reduced by a factor of 10 if the regions were interconnected. Stochastic models have long been used in the hydrological community [15,16]. Our study shows that it can also be useful for wind risk assessment in India and very likely elsewhere.

**Acknowledgement**

Funding: The authors acknowledge the financial support given by the Earth System Science Organization, Ministry of Earth Sciences, Government of India, (Grant no: IITM/MM-II/Imperial_College/2018/INT-6) to conduct this research under Monsoon Mission. We thank National Institute of Wind Energy (NIWE), Chennai for providing wind speed observation data from their monitoring stations across India.

**References:**


[1] Ministry of New and Renewable Energy, Annual Report 2018-19, 2019, https://mnre.gov.in/img/documents/uploads/file_f-1608040317211.pdf

[2] Guttikunda SK and Jawahar P, Atmospheric emissions and pollution from the coal-fired thermal power plants in India, Atmospheric Environment, 2014, 92:449-460, http://www.indiaairquality.info/wp-content/uploads/docs/2014-08-AE-Emissions-Health-Coal-PPs-India.pdf

[3] Chaturvedi R. K., Gangopadhyay A., Seshadri A., and Hiremath M., Co-benefits of power sector decarbonisation for air quality and human health in India, Policy Brief, Divecha Centre for Climate Change, Indian Institute of Science, Bengaluru, 2018, http://dccc.iisc.ac.in/assets/pdf/Policy_Brief_January_2018.pdf

[4] Staffell I. and Pfenninger S., The increasing impact of weather on electricity supply and demand, Energy, Volume 145, 15 February 2018, Pages 65-78, https://doi.org/10.1016/j.energy.2017.12.051

[5] Brayshaw DJ, Troccoli A, Fordham R and Methven J., The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: a case study over the UK. Renewable Energy, 2011, 36:2087-96, https://doi.org/10.1016/j.renene.2011.01.025

[6] Sparks N J., Hardwick S. R., Schmid M. and Toumi R., IMAGE: a multivariate multi-site stochastic weather generator for European weather and climate, Stochastic Environmental Research and Risk Assessment volume 32, pages771–784, 2018, https://link.springer.com/article/10.1007/s00477-017-1433-9

[7] Parlange M. B. and Katz R. W., An extended version of the Richardson model for simulating daily weather variables. Journal of Applied Meteorology, 2000, 39:610-622, https://doi.org/10.1175/1520-0450-39.5.610

[8] Higham, N., Computing a Nearest Symmetric Positive Semidefinite Matrix. Linear Algebra and Its Applications, 1988, 103, 103-118, http://dx.doi.org/10.1016/0024-3795(88)90223-6

[9] ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate . Copernicus Climate Change Service Climate Data Store (CDS), 2020, https://cds.climate.copernicus.eu/cdsapp#!/home

[10] Wind-turbine-models.com. Suzlon S.88-2100 power curve, 2020, available at https://en.wind-turbine-models.com/turbines/95-suzlon-s.88-2100#powercurve; accessed on 23 Nov. 20

[11] Thomas S. R., Nicolau S., Martínez-Alvarado O., Drew D. J., Bloomfield H. C., How well do atmospheric reanalyses reproduce observed winds in coastal regions of Mexico?, Meteorol Appl., John Wiley & Sons Ltd, 2021, https://doi.org/10.1002/met.2023

[12] Molina, MO, Gutiérrez, C, Sánchez, E., Comparison of ERA5 surface wind speed climatologies over Europe with observations from the HadISD dataset. Int J Climatol. 41: 4864– 4878. 2021, https://doi.org/10.1002/joc.7103

[13] Belmonte Rivas, M. and Stoffelen, A. (2019), Characterizing ERA-Interim and ERA5 surface wind biases using ASCAT, Ocean Sci., 15, 831–852, 2019, https://doi.org/10.5194/os-15-831-2019

[14] Mathew S., Wind Energy, Fundamentals, Resource Analysis and Economics, Spinger, 2016, ISBN-10 3-540-30905-5, https://www.dolcera.com/wiki/images/Wind_power_energy.pdf

463 [15] Cowden JR, Watkins DW, Mihelcic JR, Stochastic rainfall modeling in West Africa:
464 Parsimonious approaches for domestic rainwater harvesting assessment. Journal of
465 Hydrology, 2008, 361(1–2):64–77. doi:10.1016/j.jhydrol.2008.07.025
466 [16] Supit I, van Diepen CA, de Wit AJW, Wolf J, Kabat P, Baruth B, Ludwig F,
467 Assessing climate change effects on European crop yields using the Crop Growth
468 Monitoring System and a weather generator. Agriculture and Forest Meteorology,
469 2012, 164: 96–111, doi:10.1016/j.agrformet.2012.05.005
470
471
472

**Table 1.** Impact of aggregation of simulated wind generation from different plants on the fraction of days with "no generation" in IMAGE simulations of 1000 years; **a)** Set 1: plants located in Box A; Moderate improvement in no generation days; **b)** Set 2: plants located in Box B; Moderate improvement in no generation days; **c)** Set 3: Plants located in grids with maximum mean wind speed in Box A and B; **d)** Set 4: Plants located in grids with second highest mean wind speed in Box A and B; In the last two cases, improvement is evident because low-wind days in each of the two regions coincide less frequently.

| Sets | Percentage of days with "no generation" in 1000 years | | |
|---|---|---|---|
| | Plant 1 | Plant 2 | Combination |
| a) Set 1 | 10% | 14% | 9% |
| b) Set 2 | 11% | 8% | 5% |
| c) Set 3 | 10% | 8% | 1% |
| d) Set 4 | 14% | 11% | 2% |

**Figure legends**

**Figure 1.** Mean climatological wind speed at 100m above ground over India; Purple dots show 40 locations chosen for validation from wind rich regions; These 40 locations are used for ERA5 comparison with in-situ data and IMAGE validation; The green triangles indicate three locations out of these 40 locations that are chosen for validating seasonal pattern simulated by IMAGE; Brown boxes indicate the areas chosen for validation of return period of various wind speed from the simulation with respect to ERA5 input; These same regions are used for demonstrating the application of IMAGE.

**Figure 2.** Comparison of climatology ERA5 wind speed with observation data from NIWE for year 2014; A) Comparison of daily mean wind speed for 40 NIWE wind monitoring stations in wind rich regions of India; B) Comparison of monthly wind speed pattern for a NIWE wind monitoring station, Devereddypalli; The red dot indicates the monthly median wind speed from 41 years of ERA5, the vertical line shows the distribution; Although, ERA5 underestimates the wind speed, the seasonal pattern is captured accurately.

**Figure 3.A)** Comparison of mean wind speed simulated by IMAGE with ERA5 input dataset; the R square value is 0.99, and the bias is small; Each point indicate median values for an individual location. The vertical bars show the distribution across 100 ensemble members, each being as long as the input data (41 years); The x = y line is shown in black; **B)** Comparison of wind-speed spatial correlation between each of 40 locations and Dhanuskodi as simulated by IMAGE, with corresponding correlations from ERA5 input dataset; the R square value is 0.99; Each point indicates median values across 100 ensemble members, while vertical bars show the distribution; The x = y line is shown in black. **c)** Comparison of monthly average wind speed pattern simulated by IMAGE with corresponding values from the ERA5 input dataset at three high-wind locations: Dhanuskodi, Devereddypalli and Bassi.

**Figure 4.** Comparison of the probability distribution and best-fit Weibull curves based on ERA5 and IMAGE, **A)** Gudaparihar (Weibull shape factor = 3.19, scale factor = 5.31); Based on IMAGE simulated daily wind speed for 4100 years (4100 x 365 data points); **B)** Gudaparihar (Weibull shape factor = 3.19, scale factor = 5.31); Based on ERA5 based 41 years of daily wind speed **C)** Bassi (Weibull shape factor = 3.04, scale factor = 4.97); Based on IMAGE simulated daily wind speed for 4100 years (4100 x 365 data points); **D)** Bassi (Weibull shape factor = 3.04, scale factor = 4.97); Based on ERA5 based 41 years of daily wind speed.

**Figure 5. A)** Return period (x-axis) of daily average wind speed (y-axis) over Rajasthan from ERA5 reanalysis and IMAGE simulations; **B)** Return period (x-axis) of daily average wind speed (y-axis) over South India from ERA5 and IMAGE simulation; The points indicate the median return periods while the horizontal lines show the distribution of return periods.

**Figure 6. A)** Fraction of days wind speed is below 3m/s in individual pixels across India when Box A (shown in black) has a wind low. Analysis is based on 1000 years of IMAGE wind speed simulation; The blue colour represents fewer days with wind-drought; **B)** Fraction of days wind speed is below 3 m/s in individual pixels across India when Box B (shown in black) experiences a wind low.

**Figure 7.** Reduction in fraction of zero generation days in 1000 years due to aggregation of different combinations of four wind plants from different Boxes (A and B); Combination of plants from different regions demonstrates larger improvement in fraction of “no generation” days by aggregation of generation.

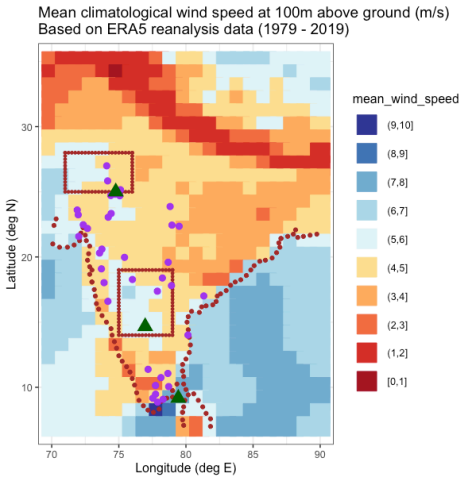



**Figure 1.** Mean climatological wind speed at 100 m above ground over India; Purple dots show 40 locations chosen for validation from wind rich regions; These 40 locations are used for ERA5 comparison with in-situ data and IMAGE validation; The green triangles indicate three locations out of these 40 locations that are chosen for validating seasonal pattern simulated by IMAGE; Brown boxes indicate the areas chosen for validation of return period of various wind speed from the simulation with respect to ERA5 input; These same regions are used for demonstrating the application of IMAGE.

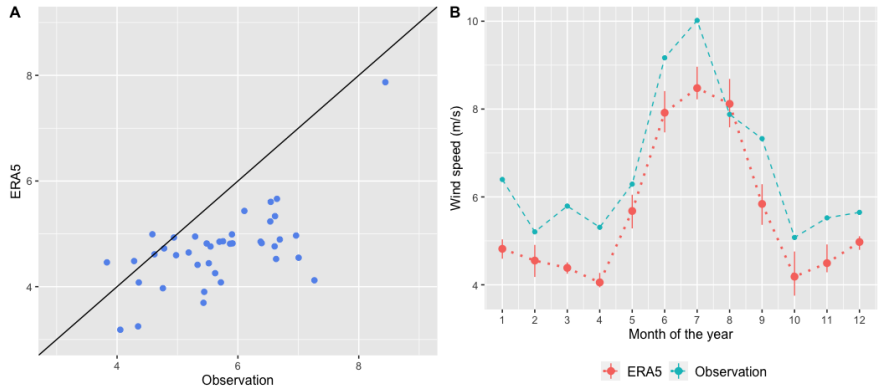



**Figure 2.** Comparison of climatology ERA5 wind speed with observation data from NIWE for year 2014; A) Comparison of daily mean wind speed for 40 NIWE wind monitoring stations in wind rich regions of India; B) Comparison of monthly wind speed pattern for a NIWE wind monitoring station, Devereddypalli; The red dot indicates the monthly median wind speed from 41 years of ERA5, the vertical line shows the distribution; Although, ERA5 underestimates the wind speed, the seasonal pattern is captured accurately.

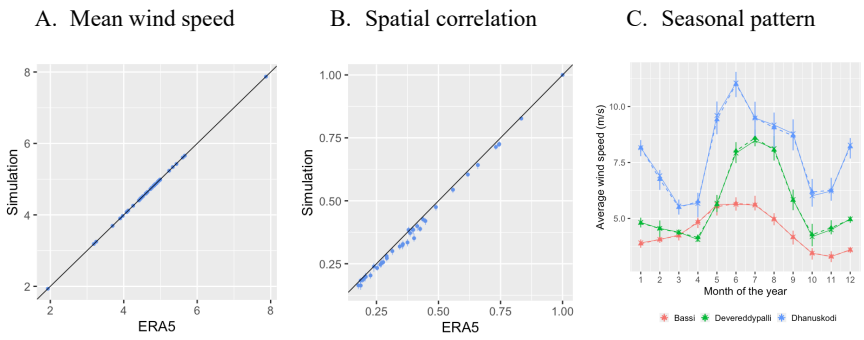
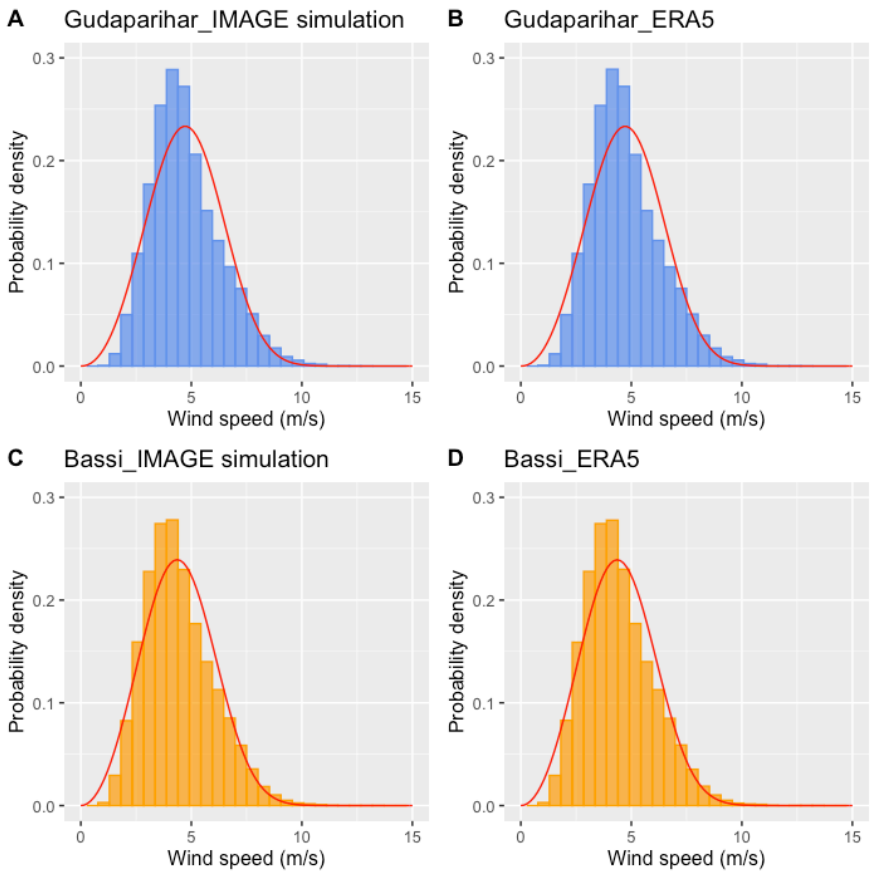



**Figure 3.A)** Comparison of mean wind speed simulated by IMAGE with ERA5 input dataset; the R square value is 0.99, and the bias is small; Each point indicate median values for an individual location. The vertical bars show the distribution across 100 ensemble members, each being as long as the input data (41 years); The x = y line is shown in black; **B)** Comparison of wind-speed spatial correlation between each of 40 locations and Dhanuskodi as simulated by IMAGE, with corresponding correlations from ERA5 input dataset; the R square value is 0.99; Each point indicates median values across 100 ensemble members, while vertical bars show the distribution; The x = y line is shown in black. **C)** Comparison of monthly average wind speed pattern simulated by IMAGE with corresponding values from the ERA5 input dataset at three high-wind locations: Dhanuskodi, Devereddypalli and Bassi.

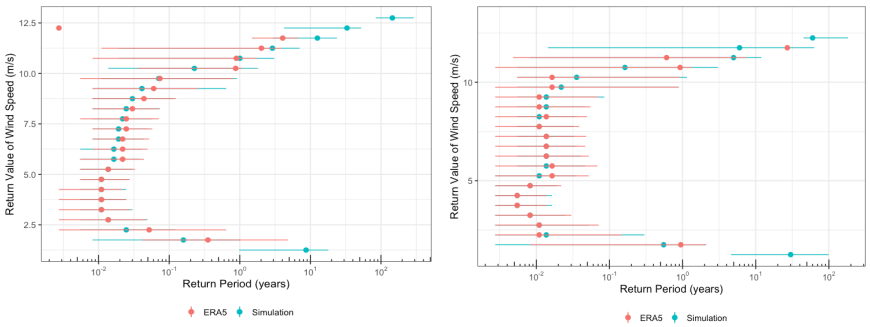



562

563 **Figure 4.** Comparison of the probability distribution and best-fit Weibull curves based on
564 ERA5 and IMAGE, **A)** Gudaparihar (Weibull shape factor = 3.19, scale factor = 5.31); Based
565 on IMAGE simulated daily wind speed for 4100 years (4100 x 365 data points); **B)**
566 Gudaparihar (Weibull shape factor = 3.19, scale factor = 5.31); Based on ERA5 based 41
567 years of daily wind speed **C)** Bassi (Weibull shape factor = 3.04, scale factor = 4.97); Based
568 on IMAGE simulated daily wind speed for 4100 years (4100 x 365 data points); **D)** Bassi
569 (Weibull shape factor = 3.04, scale factor = 4.97); Based on ERA5 based 41 years of daily
570 wind speed.

A) Rajasthan Area

B) South India

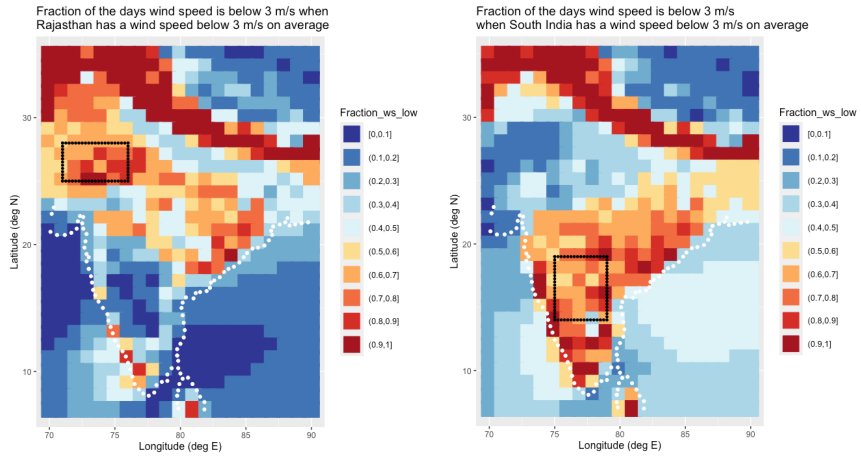



571

572 **Figure 5. A)** Return period (x-axis) of daily average wind speed (y-axis) over Rajasthan from
573 ERA5 reanalysis and IMAGE simulations; **B)** Return period (x-axis) of daily average wind
574 speed (y-axis) over South India from ERA5 and IMAGE simulation; The points indicate the
575 median return periods while the horizontal lines show the distribution of return periods.

576

A) B)

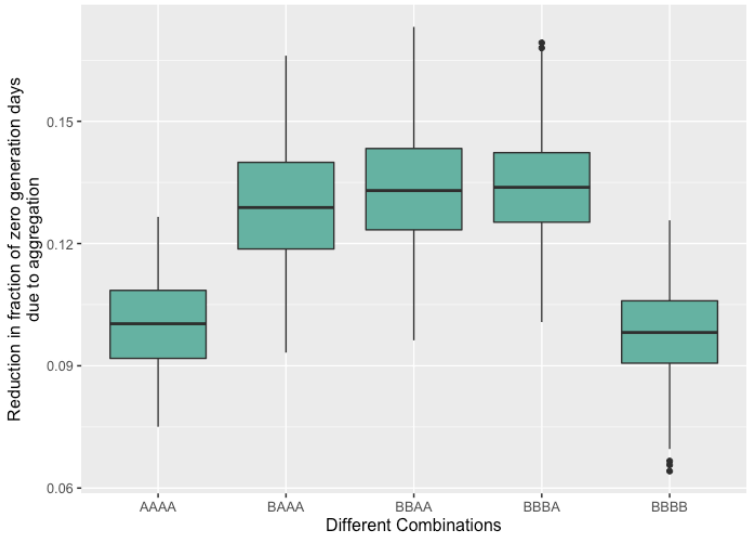



577

578 **Figure 6. A)** Fraction of days wind speed is below 3m/s in individual pixels across India
579 when Box A (shown in black) has a wind low. Analysis is based on 1000 years of IMAGE
580 wind speed simulation; The blue colour represents fewer days with wind-drought; **B)**
581 Fraction of days wind speed is below 3 m/s in individual pixels across India when Box B
582 (shown in black) experiences a wind low.

583

584




585

586 **Figure 7.** Reduction in fraction of zero generation days in 1000 years due to aggregation of
587 different combinations of four wind plants from different Boxes (A and B); Combination of
588 plants from different regions demonstrates larger reduction in fraction of "no generation"
589 days by aggregation of generation.

590

591

592