

Risk assessment of wind droughts over India

A. Gangopadhyay^{1,*}, N. J. Sparks^{2,3}, R. Toumi^{2,3} and A. K. Seshadri⁴

¹Divecha Centre for Climate Change, Indian Institute of Science, Bengaluru 560 012, India

²Blackett Laboratory, Department of Physics, Imperial College, London SW7 2AZ, United Kingdom

³Department of Physics and Grantham Institute – Climate Change and Environment, Imperial College, Exhibition Rd, South Kensington, London SW7 2BU, United Kingdom

⁴Divecha Centre for Climate Change and Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, Bengaluru 560 012, India

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 employed 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 the country.

Keywords: Decarbonization, grid interconnections, risk assessment, stochastic weather generators, wind drought.

AS part of its progress toward decarbonization, India has plans to increase the share of renewable energy (wind and solar) in its electricity system¹. 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 system⁴. 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 generation as a result of weather variability over different timescales is important for reliable operation of power systems⁴.

Wind-speed variability often arises from large-scale weather patterns⁵, leading to correlated extreme excesses or deficits in wind generation. The statistics of such extreme events typically cannot be estimated from instrumental records, because their relatively short duration does not contain many realizations of such events. Nevertheless, such extreme weather events can be the result of an underlying structure, such as the covariance in wind-

speed between different pairs of locations. Stochastic weather generators can help identify such extremes, especially when they arise from statistics estimated from shorter instrumental datasets.

Sparks *et al.*⁶ developed 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 spatio-temporal 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 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 days⁶. It uses multivariate 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 multivariate autoregressive models⁷. Here we apply IMAGE to simulate data from outside Europe and explore the risk of wind droughts across India.

IMAGE model description

We use an improved version of the IMAGE model developed by Sparks *et al.*⁶. We include only a brief description of the model here, presenting a more detailed explanation of the modifications.

*For correspondence. (e-mail: anasuya.g.research@gmail.com)

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

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

is fitted separately to each month of input data for each variable at each site, where c_s is a constant, α_s the memory parameter and ε_s is the noise term. These three parameters are each, in turn, modelled as latent Gaussian variables and 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 α_s for each month by sampling from a multivariate normal distribution. This process requires decomposition of the covariance matrix of the autoregressive parameter, Σ , to a matrix C such that $CC^T = \Sigma$. Sparks *et al.*⁶ achieved this using empirical orthogonal function decomposition. In this study we instead use Cholesky decomposition, which produces the same results but is computationally faster. In general, Σ 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 Σ is computed using the method of Higham⁸. Parameters are generated simultaneously for all 12 months in one simulated year, such that correlations between months in the same year are accurately simulated, as well as the spatial correlation between sites. The noise terms ε_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 distribution using an inverse normal quantile transformation. The pairwise Pearson's correlation coefficients 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.*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, ε_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 ε_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.

Data used and methodology

The analysis in this study was based on horizontal wind speed at 100 m over the surface from ERA5 reanalysis products by the European Centre for Medium-Range Weather Forecasts (ERA5) for 41 years (1979–2019) over the Indian region⁹. We chose 100 m above the surface because the hub heights of various onshore wind turbines lie roughly at that level. The spatial resolution of the reanalysis dataset was $0.25^\circ \times 0.25^\circ$, and temporal resolution was hourly. *In situ*, hourly 100 m wind measurement data of 40 weather stations from the National Institute of Wind Energy (NIWE), Chennai, were compared with ERA5 data to check if ERA5 could capture the overall wind speed pattern.

The methodology used was as follows. First, we tested if ERA5 could capture the overall pattern of observed wind speed over India (from NIWE measurements for 40 locations; Figure 1, purple dots). However, the aim of this study is to highlight the ability 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 dataset available to us. The input data can change, but the tool presented here remains applicable to a variety of weather variables, depending on the context.

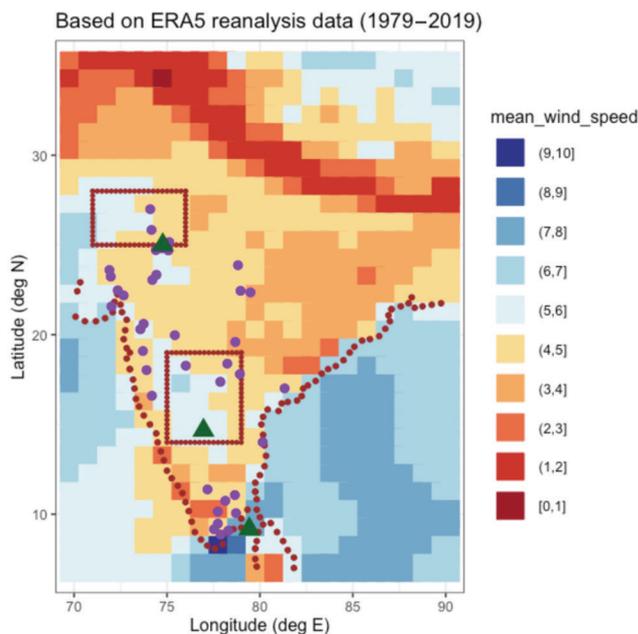


Figure 1. Mean climatological wind speed at 100 m above ground over India. Purple dots show 40 locations chosen for validation from wind-rich regions. These locations are used for ERA5 comparison with *in situ* data and IMAGE validation. Green triangles indicate three out of these 40 locations that are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

Next, the IMAGE weather generator was validated for wind speeds over the Indian region. The model was simulated for 40 locations over wind-rich regions in western and southern India for daily wind speed of 4100 years (Figure 1; purple dots). The output of 4100 years of IMAGE simulation for each of these locations was segregated into 100 ensemble members, each of 41 years in length (i.e. same length as the input data). For each of these ensemble member, the 41 years' daily wind speed time series was compared with the corresponding time series in the ERA5 input dataset. The parameters chosen for comparison were yearly mean of daily wind speed, seasonal variation of monthly mean wind speed and spatial correlation of wind speed. Three locations were used to illustrate the validation of seasonal patterns, from three different high wind-resource regions in India (Figure 1, green triangles): Dhanushkodi in the offshore region near South India, and two onshore locations in box A (Rajasthan) and box B (South India). We also examined if the wind speed simulated by IMAGE could preserve the Weibull distribution of wind speed.

As part of validation of the simulations from IMAGE, we also examined whether they could capture the low probability extreme events of high and low wind speed. This analysis compares the return periods of various wind-speed events from the ERA5 reanalysis with IMAGE simulations. The validation was done for two high wind-resource regions within Rajasthan (box A) and South India (box B) (Figure 1, brown squares).

Following the validation as described above, IMAGE was used to simulate 1000 years of wind speed over India based on the ERA5 reanalysis as input data. The ERA5 data were upscaled to $1^\circ \times 1^\circ$ resolution prior to using it as an input to IMAGE. Based on the IMAGE simulations, we estimated the probability of low wind (wind drought) over different regions and all over India. We estimated the fraction of days for which the average wind speed was below 3 m/s all over India, given that one of the wind-rich regions – box A (Rajasthan) or box B (South India) – had an average daily wind speed below 3 m/s. The 3 m/s threshold was considered because 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, with a case study, we have illustrated the benefits of having grid-connected wind plants located in different regions compared to the absence of any interconnection between regional grids. These benefits have been assessed from the perspective of 'wind drought' or 'no generation' days based on the IMAGE simulation of 1000 years. We simulated wind turbines in four sets of locations from boxes A and 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 the second highest mean wind speed in both the boxes.

We simulated one 2.1 MW wind turbine at each location. The manufacturer's power curve of the Suzlon S.88-2100 model turbine¹⁰ with a rated power output of 2.1 MW was used for converting wind speed to wind generation. A look-up table created based on the normalized power curve was used for converting the IMAGE-simulated wind speed to wind generation. Due to normalization, the rated generation from the turbine is indicated as 1. Hence, wind generation for any given wind speed can be interpreted as a fraction of the rated generation obtained at that wind speed.

To compare the benefits of having wind farms located in different sets of locations, we estimated the fraction of days in 1000 years for which there was no generation from the individual plants as well as no generation from a 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 assessed the robustness of this parameter (fraction of 'no generation' days). Twenty grids were selected from boxes A and B, which had higher average daily wind speed compared to the other grids. An equal number of grids (10) were selected from each box. We simulated all possible combinations of four grids that could be selected from among these 20 grids. One wind turbine was simulated at each location. We estimated the reduction in percentage of 'no generation' days in the aggregate generation for each combination. The reduction values were estimated based on eq. (2).

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

where f_i and f_c are the 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.

Model validation

ERA5 reanalysis dataset and observation

Before using ERA5 as input for IMAGE, we checked if it could capture wind speed over India. Figure 2 *a* compares the daily mean wind speed for 40 NIWE wind-monitoring stations in the wind-rich regions of India with ERA5. There is some underestimation of point observations of

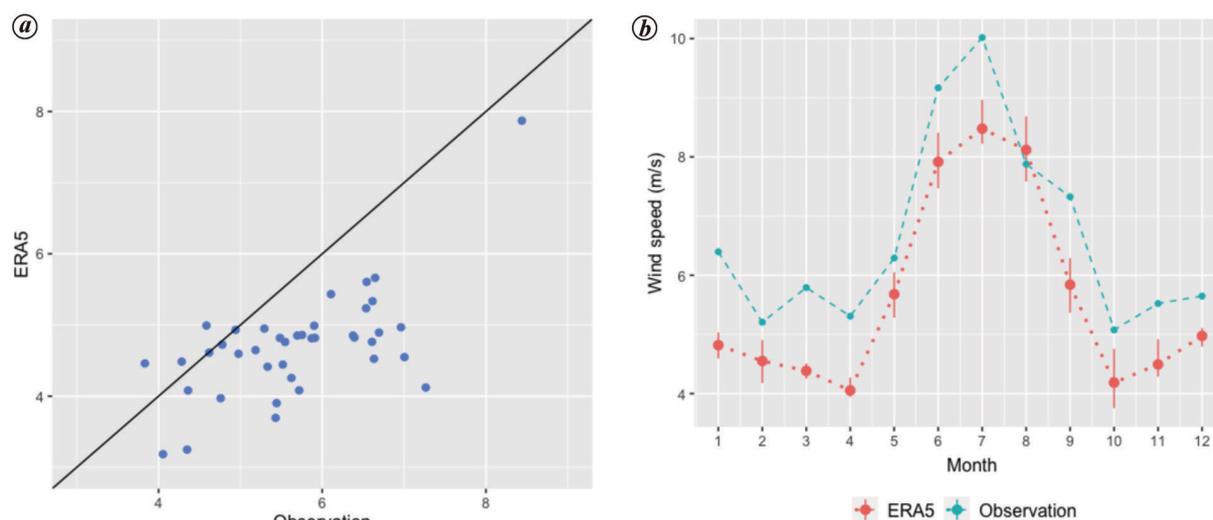


Figure 2. Comparison of climatology (average (1979–2019)) ERA5 wind speed with observation data from NIWE for the year 2014. *a*, Comparison of daily mean wind speed for 40 NIWE wind monitoring stations in wind-rich regions of India. *b*, Comparison of monthly wind speed pattern for a NIWE wind monitoring station, Devereddyapalli, Andhra Pradesh. Red dot indicates the monthly median wind speed from 41 years of ERA5 and vertical line shows the distribution. Although ERA5 underestimates wind speed, the seasonal pattern is captured accurately.

wind speed in ERA5, which might result from spatial averaging of wind speed in the gridded dataset. ERA5 captured the spatial pattern of daily mean wind speed with moderate accuracy (correlation coefficient of 0.58 with P -value ≈ 0). However, it is important to note that the aim of this study is to highlight the ability 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 has outperformed ERA-Interim and MERRA-2 consistently across several parts of the world and so this reanalysis has been recommended for local wind-power studies¹¹. Molina *et al.*¹² reported that ERA5 could reproduce the wind-speed spectrum over Europe. The authors noted that despite shortcomings, ERA5 provided a regular spatial and temporal wind distribution that is important for renewable energy studies¹². Belmonte Rivas and Stofelen¹³ analysed the differences between ERA-Interim and ERA5 surface wind fields relative to Advanced Scatterometer (ASCAT) ocean vector wind observations, and found that ERA5 performed better than ERA-Interim in terms of mean and transient wind errors.

Figure 2 *b* compares the observed and ERA5 monthly wind-speed patterns for a NIWE wind-monitoring station, Devereddyapalli, Andhra Pradesh. The red dot indicates monthly median wind-speed from 41 years of ERA5, while the vertical line shows the distribution. Although ERA5 underestimates wind speed, the seasonal pattern is captured accurately.

Mean wind speed, seasonal pattern and spatial correlation

Simulations from IMAGE are able to well represent the mean wind speed over India. The validation of IMAGE

simulations was 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; the bias was small and R^2 value was 0.99. This comparison is made in Figure 3 *a*, where the dots show the median values and vertical lines indicate the distribution of the ensemble members for each location. This distribution is narrow, indicating that each of the ensemble members closely approximates the mean wind speed.

Next, we assessed if IMAGE simulations could represent the spatial correlation of wind speed between various locations. We considered Dhanushkodi in Tamil Nadu as a reference location, and compared the spatial correlation coefficient between daily wind speed for ERA5 input data and IMAGE simulations between Dhanushkodi and each of the other 40 locations (Figure 3 *b*). Each of the dots in Figure 3 *b* indicates the median of the 100 ensemble members from the simulation plotted against the ERA5 value, between Dhanushkodi and another 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^2 value of a linear regression between the correlation coefficients of input and simulated output was 0.99, and the bias was small.

Much of India experiences high wind speed during the summer monsoon months (JJAS – June, July, August and September). Figure 3 *c* shows the seasonal variation of wind speed at three locations indicated in Figure 1: Dhanushkodi, Devereddyapalli and Bassi. For each location, the solid lines indicate the ERA5 inputs while the dashed lines indicate distributions from the IMAGE simulations.

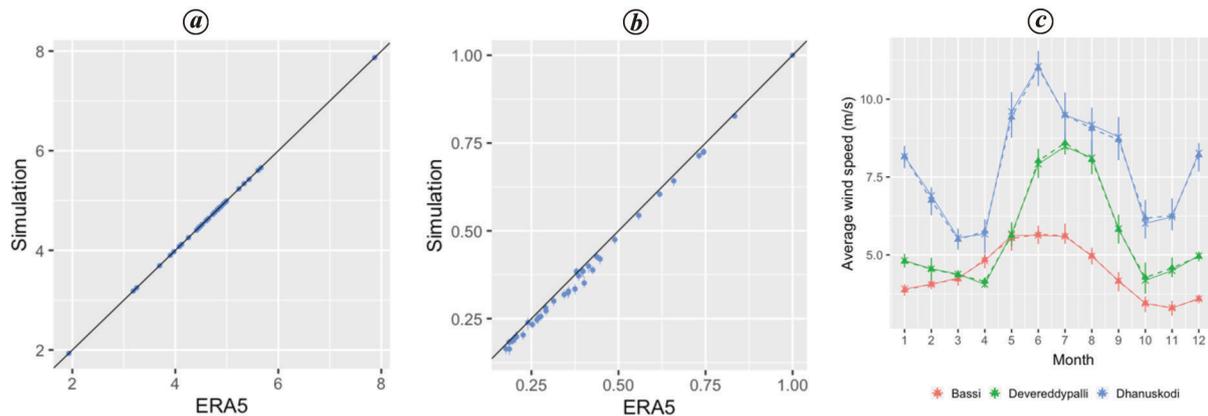


Figure 3. *a*, Comparison of mean wind speed simulated by IMAGE with ERA5 input dataset the R^2 value is 0.99 and bias is small. Each point indicates median values for an individual location. The vertical bars show distribution across 100 ensemble members, each being as long as the input data (41 years). The $x = y$ line is shown in black. *b*, Comparison of wind speed spatial correlation between each of 40 locations and Dhanushkodi, Tamil Nadu, as simulated by IMAGE, with corresponding correlations from ERA5 input dataset; the R^2 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: Dhanushkodi, Deverreddypalli, Andhra Pradesh and Bassi, Rajasthan.

The crosses and 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.

Wind speed distribution pattern

Wind speeds are expected to follow a Weibull distribution¹⁴. Figure 4 compares the probability distribution plot of the ERA5 input data for 41 years and IMAGE-simulated wind speed for 4100 years (4100×365 data points) for two locations, viz. Gudaparihar and Bassi. The red lines indicate that the Weibull distribution fits the appropriate shape and scale factor for the distributions. This shows that the IMAGE simulations could preserve the wind-speed distribution.

Return periods

We estimated the return periods of different values of wind speed for both the wind-rich regions (boxes A and B) shown in Figure 1. Figure 5 *a* and *b* shows 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, such return periods can also be

estimated from the IMAGE simulation. A similar result is observed in 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 5 *a* 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 has been 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 the ERA5 dataset and the median return period of this event was predicted as 60 years by IMAGE. For box B, IMAGE can simulate an extremely 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 model.

Results

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 two distinct wind-rich regions in the country. Since wind development is likely to concentrate substantially in these regions, we examined the association between low-wind days in both of the regions. IMAGE simulations over a period of 1000 years indicate that, on days when box A (Rajasthan) experiences low wind on average (below 3 m/s), there is only 0–10% probability that locations in box B (South

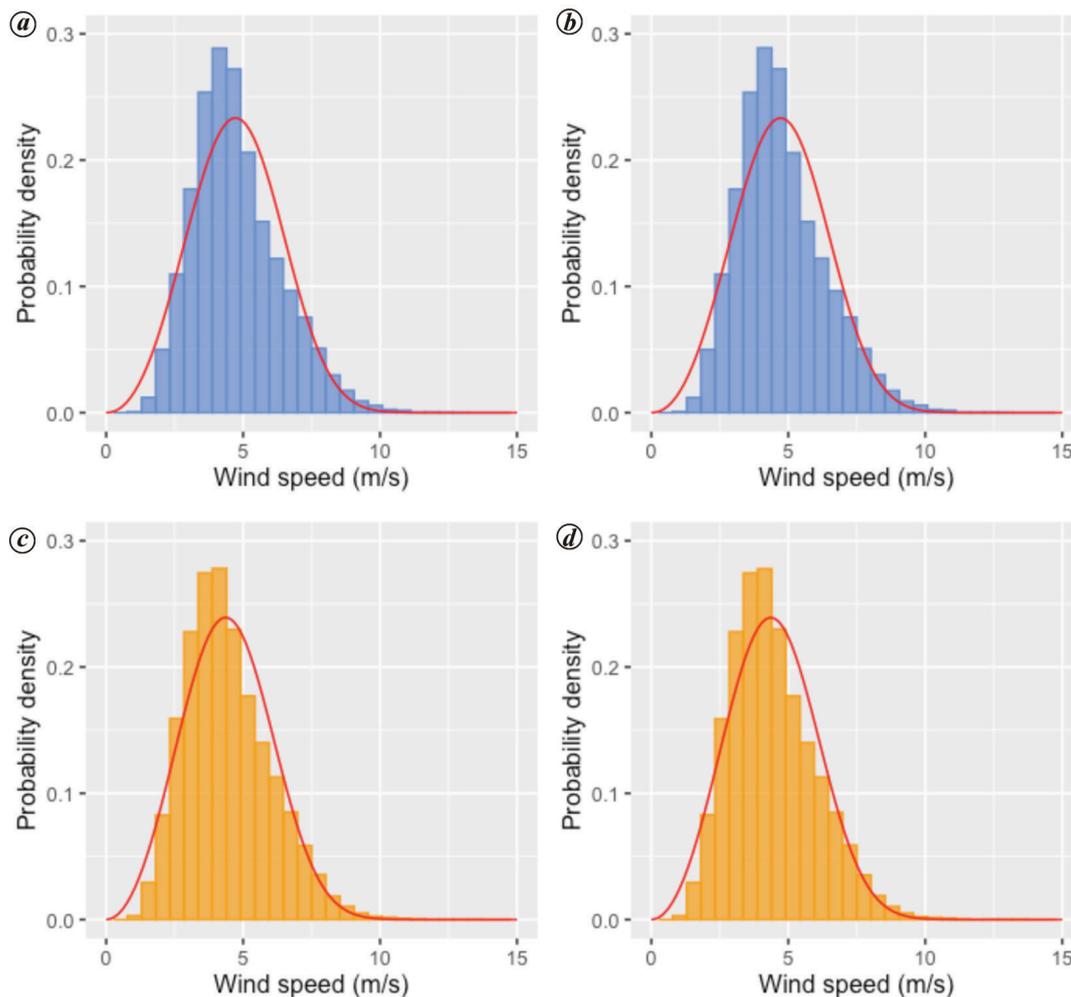


Figure 4. Comparison of probability distribution and best-fit Weibull curves based on ERA5 and IMAGE. *a*, Gudaparihar (Weibull shape factor = 3.19, scale factor = 5.31), based on IMAGE-simulated daily wind speed for 4100 years (4100×365 data points). *b*, Gudaparihar, Madhya Pradesh (Weibull shape factor = 3.19, scale factor = 5.31) from 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 (4100×365 data points). *d*, Bassi (Weibull shape factor = 3.04, scale factor = 4.97); from ERA5-based 41 years of daily wind speed.

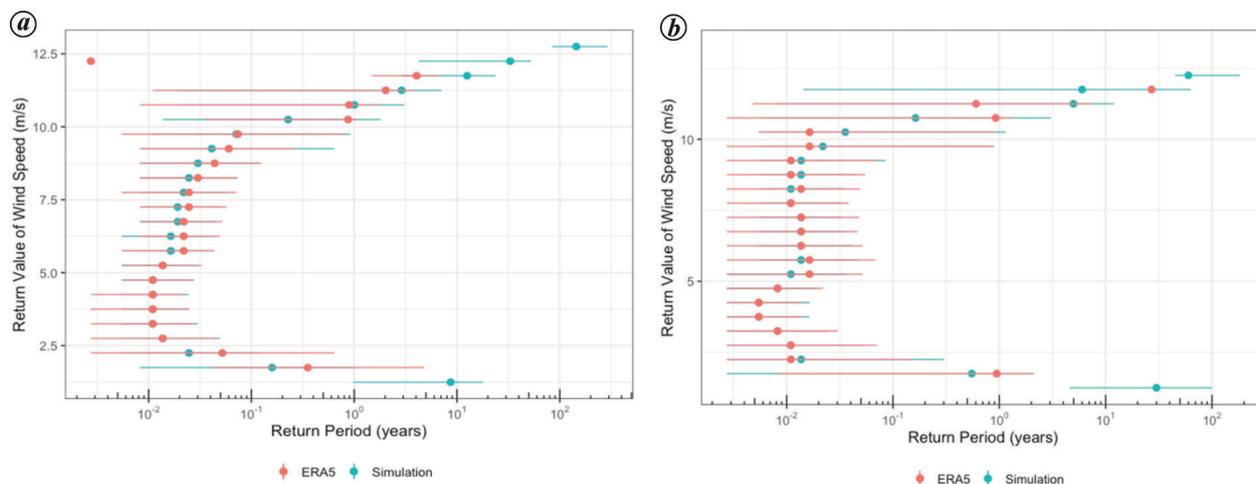


Figure 5. Return period (*x*-axis) of daily average wind speed (*y*-axis) over (a) Rajasthan and (b) South India from ERA5 reanalysis and IMAGE simulations. The points indicate the median return periods, while horizontal lines show the distribution of return periods.

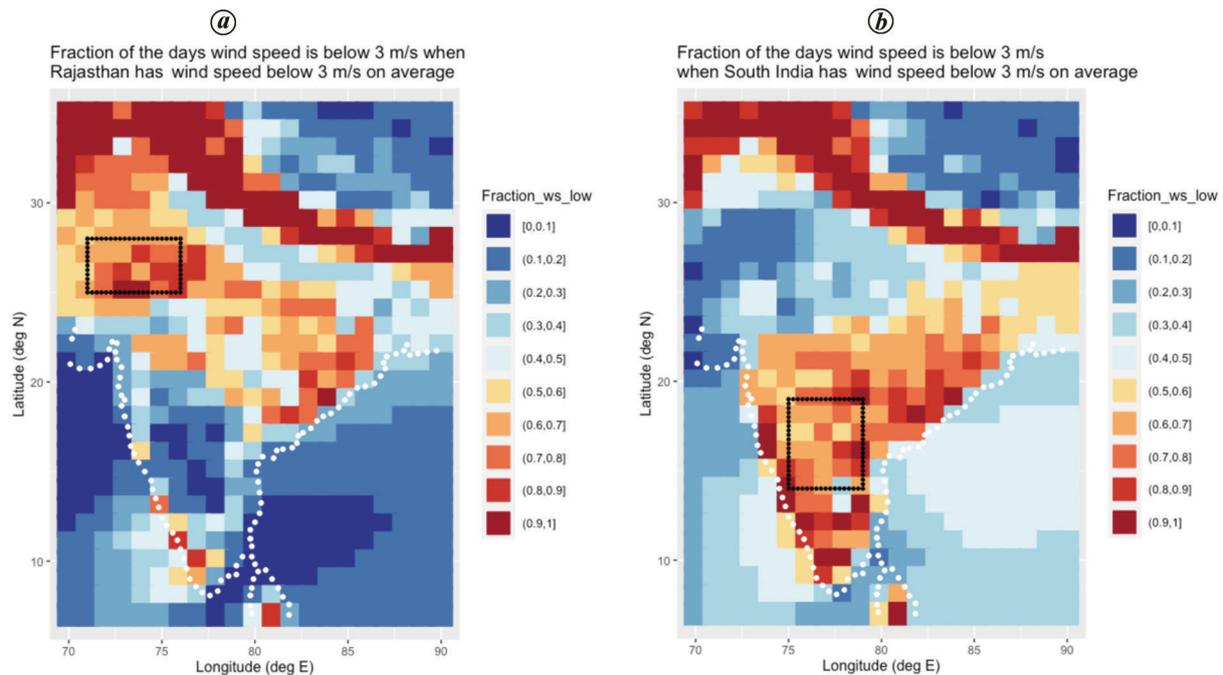


Figure 6. *a*, Fraction of days when wind speed is below 3 m/s in individual pixels across India when box A (shown in black) experiences a wind low. Analysis is based on 1000 years of IMAGE wind speed simulation. Blue colour represents fewer days with wind drought. *b*, Fraction of days when wind speed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

India) will have winds lower than 3 m/s (Figure 6 *a*). Low wind in box A occurs in 7% of the days in 1000 years. Similarly, on the days when box B has low wind, with average below 3 m/s, there is only 10%–20% probability that individual locations within box A will also have low wind (Figure 6 *b*). Low wind in box B occurs 6% of the time in 1000 years. Clearly, these two regions demonstrate a complementary behaviour from the perspective of wind drought. Hence, there is a possibility of avoiding grid-wide wind droughts if regional grids in these two regions are themselves connected. This is illustrated further with the help of a case study.

Case study

Four sets of paired locations from box A (Rajasthan) and box B (South India) were chosen for the case study. As both boxes A and B demonstrate a complementary behaviour from the perspective of wind drought, these sets of pairs were chosen accordingly. Daily generation from one wind turbine was simulated for each location depending on the local wind speed.

- 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 each of the boxes.

Set 4: Grids with the second highest mean wind speed in each of the boxes.

We estimated the fraction of days with zero generation in the 1000 years IMAGE simulations, for the individual plants and a combination of the two plants, one in each of the chosen grids. In the case of set 1, the individual plants in Rajasthan had 11% and 15% of the days with zero generation respectively, while the combination of these two plants had 10% of days with zero generation. Similarly, in the case of set 2, the individual plants in South India had 11% and 9% of the days with zero generation respectively, whereas the combination of these two plants had 5% of days with zero generation (Table 1). These two case studies demonstrate that in the 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 estimated the fraction of days with zero generation in the 1000 years IMAGE simulations for each location as well as their combination. We found that

while the plants in boxes A and B experienced 11% and 9% of the days with zero generation respectively, the combination of these two plants experienced only 1% of the days with zero generation. Set 4 considers the combination of grids with the 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 had 15% and 11% of the days with zero generation respectively, their combination had only 2% of the days with zero generation (Table 1). For cases 3 and 4, improvement in the 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.

Next we examined the reduction in the fraction of ‘no generation’ days. Figure 7 depicts the results. After identifying 20 grids with the highest mean wind speeds in boxes A and B, we simulated the aggregate generation from all possible combinations of the four grids chosen from these 20 grids. The boxplot indicates the reduction in fra-

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. Set 1: Plants located in box A; moderate improvement during ‘no generation’ days. Set 2: Plants located in box B; moderate improvement during ‘no generation’ days. Set 3: Plants located in grids with maximum mean wind speed in boxes A and B. Set 4: Plants located in grids with the second highest mean wind speed in boxes A and B. In the last two cases, improvement is evident because low-wind days in each of the two regions coincide less frequently

Set	Percentage of days with ‘no generation’ in 1000 years		
	Plant 1	Plant 2	Combination
1	10	14	9
2	11	8	5
3	10	8	1
4	14	11	2

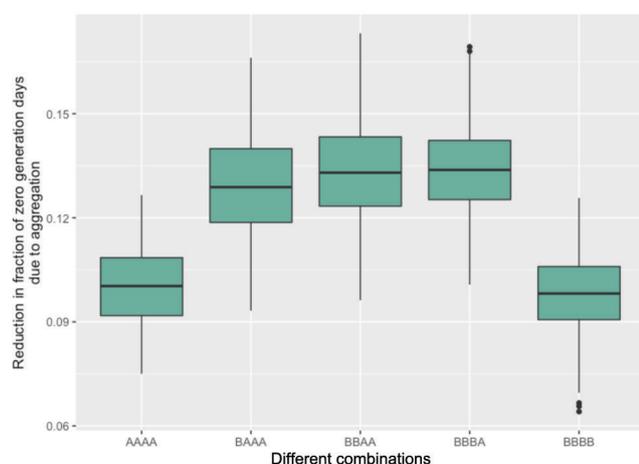


Figure 7. Reduction in fraction of zero generation days in 1000 years due to aggregation of different combinations of four wind plants from different boxes (A and B). Combination of plants from different regions demonstrates larger reduction in fraction of ‘no generation’ days by aggregation of generation.

tion 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 in Figure 7 has all four grids from box A, while the combination AABB has two grids from each box (A and 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 75th and 25th percentile values for each combination respectively. Combination of plants from different regions (AAAB, AABB, BBBA) demonstrates larger reduction (0.13–0.14 (median)) in the fraction of ‘no generation’ days by aggregation of generation compared to plants located in the same region (AAAA, BBBB) (reduction of 0.1).

Discussion and conclusion

The IMAGE weather generator has been validated for wind-power studies over India. It could reproduce the statistics of ERA5 reanalysis over the country. 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 insights into the correlated behaviour of wind droughts in the important wind-resource regions of Rajasthan and South India. Our emphasis here is on highlighting the ability of a novel tool to study renewable drought over India, using the relevant gridded datasets, and 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 years and other low-probability events. A 1000-yr simulation of daily wind speeds allows us to quantify the likelihood of wind droughts anywhere in India.

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

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

2. Guttikunda, S. K. and Jawahar, P., Atmospheric emissions and pollution from the coal-fired thermal power plants in India. *Atmos. Environ.*, 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 Pfenniger, S., The increasing impact of weather on electricity supply and demand. *Energy*, 2018, **145**, 65–78; <https://doi.org/10.1016/j.energy.2017.12.051>.
5. Brayshaw, D. J., 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. *Renew. Energy*, 2011, **36**, 2087–2096; <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 multisite stochastic weather generator for European weather and climate. *Stoch. Environ. Res. Risk Assess.*, 2018, **32**, 771–784; <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. *J. Appl. Meteorol.*, 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 Appl.*, 1988, **103**, 103–118; [https://dx.doi.org/10.1016/0024-3795\(88\)90223-6](https://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, 2020; <https://cds.climate.copernicus.eu/cdsapp#!/home>
10. Wind-turbine-models.com. Suzlon S.88-2100 power curve, 2020; <https://en.wind-turbine-models.com/turbines/95-suzlon-s.88-2100-#powercurve>
11. Thomas, S. R., Nicolau, S., Martinez-Alvarado, O., Drew, D. J. and Bloomfield, H. C., How well do atmospheric reanalyses reproduce observed winds in coastal regions of Mexico? *Meteorol. Appl.*, 2021, **28**(5); <https://doi.org/10.1002/met.2023>.
12. Molina, M. O., Gutierrez, C. and Sanchez, E., Comparison of ERA5 surface wind speed climatologies over Europe with observations from the HadISD dataset. *Int. J. Climatol.*, 2021, **41**, 4864–4878; <https://doi.org/10.1002/joc.7103>.
13. Belmonte Rivas, M. and Stoffelen, A., Characterizing ERA-interim and ERA5 surface wind biases using ASCAT. *Ocean Sci.*, 2019, **15**, 831–852; <https://doi.org/10.5194/os-15-831-2019>.
14. Mathew, S., *Wind Energy, Fundamentals, Resource Analysis and Economics*, Springer, 2016; ISBN-103-540-30905-5.
15. Cowden, J. R., Watkins, D. W. and Mihelcic, J. R., Stochastic rainfall modeling in West Africa: parsimonious approaches for domestic rainwater harvesting assessment. *J. Hydrol.*, 2008, **361**(1–2), 64–77; doi:10.1016/j.jhydrol.2008.07.025.
16. Supit, I., van Diepen, C. A., de Wit, A. J. W., Wolf, J., Kabat, P., Baruth, B. and Ludwig, F., Assessing climate change effects on European crop yields using the crop growth monitoring system and a weather generator. *Agric. For. Meteorol.*, 2012, **164**, 96–111; doi:10.1016/j.agrformet.2012.05.005.

ACKNOWLEDGEMENTS. We acknowledge financial support from the Earth System Science Organization, Ministry of Earth Sciences, Government of India (grant no: IITM/MMII/Imperial_College/2018/INT-6) to conduct this study under Monsoon Mission. We thank the NIWE, Chennai for providing wind speed observation data from their monitoring stations across India.

Received 5 April 2021; revised accepted 1 March 2022

doi: 10.18520/cs/v122/i10/1145-1153