

Synergising Simulated Annealing and Generative Adversarial Network for Enhanced Wind Data Imputation in Climate Change Modelling

Soumyabrata Bhattacharjee* and Gaurav Kumar Gugliani

Mandsaur University, Mandsaur, MP, India

✉ soumyabrata.bhattacharjee@meu.edu.in

Received October 2, 2023; revised and accepted October 19, 2023

Abstract: Climate models help us simulate and predict how the Earth's climate is going to change in the future. Wind speed data is critical for developing and validating such models. However, in the real world, often owing to many factors such as station maintenance and sensor failures, a considerable amount of wind data goes missing. The Generative Adversarial Network (GAN) has been used to impute missing wind data, but the handling of unrealistic GAN output has remained largely unstudied. In this paper, we propose a novel hybrid approach that combines both the GAN and dual annealing algorithms to not only impute missing wind speed data but also counter unrealistic GAN outcomes. The hourly mean wind data has been collected from the National Centers for Environmental Information for four Indian stations, viz. Ahmedabad, Indore, Mangaluru and Mumbai. We compared the performance of the proposed approach with those of k-nn, soft imputation, and plain GAN-based approaches on mean, variance, standard deviation, kurtosis, skewness, and R-square. We found that our approach ranks number one based on the R-square value for all the considered stations. Our model consistently produces realistic results, unlike plain GAN. We observed that Mumbai has the lowest percentage of missing data (13.14%) and the highest R-square value (0.9999186451). However, Indore has the highest percentage of missing data (46.6463%) and the lowest R-square value (0.9046885604).

Keywords: Climate modelling; Wind speed data; Imputation; Generative adversarial network; Dual annealing algorithm.

Introduction

Climate models are crucial tools for comprehending and anticipating climate change (Reichle, 2023). Wind speed data is a key input for these models (Lakku & Behera, 2022), but it can have some missing values, due to various factors, such as station maintenance and sensor failures (Boomgard-Zagrodnik & Brown, 2022). This can introduce uncertainty into the models and lead to inaccurate estimations of regional climate change (Yao et al., 2023), aquatic ecosystem (Duvall et al., 2022), sea surface temperature (Wills et al., 2022),

sea level pressure (O'Reilly et al., 2021), atmospheric circulation (Stevens & Bony, 2013), wind gusts (Cheng et al., 2014), rain and storms (Voosen, 2020). Estimation of extreme climate events (Outten & Sobolowski, 2021), such as cyclones (Xi et al., 2023) and their impact on society, also get downgraded because of missing wind data (Schewe et al., 2019). The accuracy of climate models, to predict impacts of climate change on wind energy also gets hampered (Doddy Clarke et al., 2022). Hence, in recent years imputing missing wind data has gained traction, as shown in Table 1.

*Corresponding Author

Table 1: Summary of the literature published on wind data imputation since 2018

<i>Study</i>	<i>Method explored for imputation</i>	<i>Geographical location considered</i>
(Sareen et al., 2023)	k-nearest neighbor (k-nn)	Bhogat, Chandori, Charanka, Gandhi Nagar, Surat
(Kosana et al., 2022)	Generative Adversarial Network (GAN)	Idalia, Bend
(Silei et al., 2023)	Matrix Completion Process based on Singular Value Thresholding (SVT) method	Kedros
(Jha et al., 2022)	Deterministic Approach based on optimal value of speed-power constant	Arizona, Utah, Colorado
(Jing et al., 2022)	Improved Context Encoder network	Hunan
(Qu et al., 2020)	Combination of GAN and Travelling Salesman Problem (TSP)	A North Chinese wind farm (the authors did not mention anything more about the location)
(Afrifa-Yamoah et al., 2020)	Structural Time Series Model with Kalman smoothing, Autoregressive Integrated Moving Average (ARIMA) with Kalman smoothing, Multiple Regression Modelling	Esperance, Perth, Learmonth, Broome
(Tawn et al., 2020)	Inverse Probability Weighting, Mean Imputation (MI), Multiple Imputation using Markov Chain Monte Carlo method	10 wind farms located in the UK, 30 others located in other parts of Europe (the authors did not mention anything more about the location)
(Liao et al., 2022)	Context Encoder, Auto-Encoder, K-Means, k-nn, Back Propagation Neural Network, Cubic Interpolation, Conditional GAN	National Wind Technology Center (NWTC), Cape May, Butler Grade, Bovina, Megler, Cochran
(T. Liu et al., 2018)	Multiple Imputation using Mixture Model and Expectation–Maximization algorithm, Deletion Method, MI, k-nn	Jiangsu province
(Özen & Deniz, 2022)	CatBoost model	Urla wind power plant
(Ayiah-Mensah et al., 2021)	Multiple Imputation using Multivariate Imputation with Chain Equations (MICE)	Brest
(Ammar & Xydis, 2023)	Last observation carried forward (LOCF), MI	Tejona, Nørrekær Enge, Jade wind park, Wind Turbine in Turkey (the authors did not mention anything more about the location of the station from Turkey)
(Faybishenko et al., 2022)	Seasonally Decomposed Missing Value Imputation	East River Watershed (Colorado), Barro Colorado Island
(Ben et al., 2020)	Linear Regression	Tsetang, Lhasa, Nyêmo

In Table 1, the heuristic methods such as the deletion method simply delete the missing values, which may lead to bias in the analysis (Donner, 1982). Another heuristic method LOCF, replaces any missing value with the last observed value (Hamer & Simpson, 2009). LOCF is based on the assumption that the result of the analysis is unaffected by the missing data, which in the context of wind appears improbable, as wind flow is highly stochastic (Kayalvizhi & Kumar, 2018).

Statistical methods such as MI substitute all the missing values with the mean of the observed values. While it increases the sample size, it does not add any information (Malhotra, 1987). Another popular statistical method to impute missing data is Multiple Imputation. It imputes the missing data with multiple plausible values while preserving the natural variation and probability distribution of the existing values (Kang, 2013). It imputes the missing value of one variable, multiple times, till a satisfactory value has been obtained, with reference to the existing data for other variables. This very fact reduces its usability in imputing missing wind data because the whole weather station, which records other meteorological parameters too, may stop recording data briefly for different reasons (Yassen & Abdul Kareem, 2023). Besides, this method is very slow (Burgette & Reiter, 2010) and computationally expensive (Junger & Ponce de Leon, 2015), as it first creates multiple datasets with plausible values for the missing data points. Then each of these datasets is statistically analysed, and a final data set is created based on this analysis.

Another statistical missing data imputation method is Inverse Probability Weighting. In this method, the probability of missing data is first determined based on the observed data. The inverse of this probability is then used as the weight. These weights are then used in imputing missing values that are adjusted to account for the missingness and maintain the underlying patterns in the observed data. However, one major drawback of this method is that it is very sensitive to the assumptions of the model used to estimate the probabilities (Sun & Tchetgen Tchetgen, 2018). A little deviation in the assumptions may lead to biased imputations.

Linear regression is a simple technique to impute the missing data statistically. Though it has been found to work well in imputing missing wind speed data (Ben et al., 2020), it has not been applied to impute missing wind direction data yet. This may be attributed to the fact that wind direction is nonlinear (Farrugia & Micallef, 2006) and linear regression performs poorly in such cases (Bertsimas et al., 2017).

Cubic interpolation is another simple statistical method for imputing missing data. It imputes the missing data, by fitting the available data points with a third-degree polynomial. It is preferred when the imputed values do not vary wildly (Liu et al., 2020). This very fact renders it ineffective in imputing missing wind data, as atmospheric wind is highly stochastic in nature.

The seasonal decomposition technique is another statistical method that has recently gained a lot of popularity in imputing missing wind data. In this method, the wind data is first decomposed into three components: the trend, the seasonal component, and the residual/noise component. Each of these components is then individually imputed, and the imputed components are combined to obtain the final values. One major problem with this approach is that it may fail to effectively capture outliers (Battaglia et al., 2020) due to the complex seasonal pattern of wind (Charakopoulos et al., 2019). Similarly, the CatBoost model, which is a hybrid of Gradient Boosting and Decision Trees algorithms (Prokhorenkova et al., 2018), also requires the elimination of outliers, before fitting the wind data (Özen & Deniz, 2022). Whereas, considering outliers in wind data is critical for ensuring the safe operation of wind farms (Ye et al., 2016).

Multiple Regression Model, which imputes missing wind speed based on other meteorological parameters, has been found to outperform hybrid methods that combine the Kalman filter with the Autoregressive Integrated Moving Average (ARIMA) method and Structural Time Series Model (Afriya-Yamoah et al., 2020). However, the major problem with this type of regression model is that the weather station may stop recording all the meteorological parameters randomly owing to sudden disruption in power supply and telecommunication network (Park & Baek, 2023).

The mathematical method for matrix completion like Singular Value Thresholding (SVT) Algorithm has also been used for imputing missing values of wind speed (Silei et al., 2023). SVT is an iterative process, which continues till satisfactory singular values are obtained for a low rank matrix, which is similar to the observed data (Cai et al., 2010). However, this method doesn't perform well, if a large amount of data goes missing (Tan et al., 2015).

Neural Networks, like Back Propagation Neural Network (BPNN), Generative Adversarial Network (GAN), Context Encoder (CE) and Auto Encoder (AE) have also been tested in imputing missing wind data. Neural networks are data-driven models. In any neural

network, there are layers of artificial neurons, which are interconnected with each other. During training, the information is passed between the input and output neural layers, through intermediate connected layer(s). The difference between the actual value and the output of the network is the error. In BPNN, this error is propagated backward through the network. The weights of the network are then adjusted iteratively, using the gradient descent method, till the error is minimised to an acceptable level (Rumelhart et al., 1986).

GAN is a combination of two neural networks: Generator and Discriminator. These two networks pit each other. The Generator network tries to produce synthetic data, which tries to mimic actual observation. Whereas, the Discriminator model tries to distinguish the synthetic data from the actual data (Goodfellow et al., 2014). When the synthetic data becomes indistinguishable from the actual data, then we can say that the model has been properly trained and can now be used for imputing the missing data. However, if the Discriminator becomes too good at identifying the synthetic data, then the Generator won't learn anything meaningful, and the GAN will produce meaningless output. This is a very common problem with GAN and is known as the vanishing gradient problem (Brophy et al., 2023).

Auto-Encoder is a kind of neural network, which consists of two neural networks: Encoder and Decoder. Unlike GAN, where two neural networks pit each other, here they work together. The observed wind data is first fed to the Encoder, which compresses the data and stores it in a latent space, called 'code'. The Decoder then takes this code and tries to reconstruct the original data, by decompressing it (Goodfellow et al., 2016). If the Decoder could reconstruct the data properly, then the Autoencoder is said to have captured the salient features of the data and learned from its compressed representation, else the whole combination needs to be retrained. Once the desired level of accuracy has been achieved, then the missing wind data may be imputed using Autoencoders.

Context-Encoder is another type of neural network, which consists of three parts: Encoder, Context Decoder and a Discriminator. The observed wind data is first fed to the Encoder network, which compresses the data and stores it in a latent space. That compressed data is then taken by the Context Decoder and the original data is then reconstructed, but this time, with an added constraint of considering the context of the data, such as its Probability Distribution Function. The performance of the Context Decoder is evaluated by the Discriminator

Network. This network tries to discriminate between the actual and generated data (Liao et al., 2022). If the discriminator cannot distinguish between the generated and the actual data, then the network may be said to be properly trained and then the model may be used to impute missing wind data.

However, the performance of Auto-Encoder, Context Encoder and GAN, in imputing missing wind speed and direction data, have been found to be similar (Liao et al., 2022).

Unlike complex and computationally intensive neural networks (Han et al., 2015), k-nn is a very simple data-driven imputation model (Shataee et al., 2012). In the k-nn imputation model, the missing data are imputed by taking the average of the k nearest neighbouring values of the missing value (Batista & Monard, 2002). This simple method has been successful in accurately imputing missing wind speed data in several places in India (Sareen et al., 2023). Whereas in places, which are in other countries, k-nn seems to perform poorly in imputing missing wind data (Liao et al., 2022; T. Liu et al., 2018).

From Table 1, we observed that there is no single universally acceptable imputation method for wind data. This may be attributed to the fact that the behaviour of wind changes from one place to another (Gugliani, 2020). Furthermore, we also observed that most of the published works focus on imputing missing wind speed data, while the research on the imputation of missing wind direction data is limited.

In this paper, we initially studied the performance of k-nn, Soft Imputing (SI), and GAN in four Indian stations for imputing missing wind speed data using Mean, Variance, Standard Deviation, kurtosis, skewness, and R-square values. Then we proposed a novel method that is a combination of GAN and the Dual Annealing Algorithm.

Geographical Conditions and Observation Period

In the present study, hourly wind speed (in m/s) has been collected from the National Centers for Environmental Information (NCEI) (*Data Search | National Centers for Environmental Information (NCEI)*, n.d.). The period of study is from January 1, 2017, at 12:00 AM to January 1, 2023, at 12:00 AM. The data are collected at a height of 10 m above ground level, in the vicinity of airports, to ensure that the wind flow is unhindered by local terrain or any other obstruction (Gugliani, 2020). In this paper, we considered four Indian stations, viz. Ahmedabad, Indore, Mangaluru and Mumbai.

The latitude, longitude, elevation relative to mean

sea level, total number of observations and percentage of missing wind data, for each of the stations are summarized in Table 2. In Table 2, the column “**Total Number of Observation**” represents the number of observations that passed all the quality checks by the issuing agency (*FEDERAL CLIMATE COMPLEX*, n.d.).

From Table 2, we see that the stations considered in the study are spread across different regions in India and the percentage of missing wind data varies with them. During the study period, apart from Mumbai, the percentage of missing data at all other places was more than 20%. Indore reports the highest percentage of missing data (46.6464 %), whereas Mumbai reports the lowest percentage of missing data (13.2015 %).

The places are chosen strategically so that the performance of the methods can be evaluated in coastal as well as inland regions. Coastal cities, Mangaluru and Mumbai, along the Arabian Sea, have been chosen as they experience strong and consistent wind from the sea (Kumar et al., 2002). The coastal cities spread from the western to the southern parts of India.

The inland cities of Ahmedabad and Indore are more diverse in their terrain. Ahmedabad is a dry, sandy and mostly flat region (Makwana & Gandhi, n.d.). On the other hand, Indore is a semi-arid region in the Malwa plateau in the central part of India (Pandey, n.d.). All the considered stations, be it coastal or inland, vary in elevation with respect to the mean sea level.

Mathematical Analysis

k-Nearest Neighbour (k-nn) Imputation Method

Mathematical expression for k-nn imputation, adopted from (Sareen et al., 2023), is shown in Equation 1.

$$\text{missing value} = \sum_i \frac{C_i W_i}{W_i}$$

Where k is the number of neighbour points, i is the number of the considered neighbour, C_i is the value of the considered neighbour and W_i is the reciprocal of the distance between the considered neighbour and the missing value.

The mathematical expression for calculating w_i , is also adopted from (Sareen et al., 2023), and is shown in Equation 2.

$$w_i = \frac{1}{d_i}$$

In Equation 2, d_i represents the Euclidian distance between the missing value and the considered neighbour. In this study, the square root of the sample size of the dataset has been chosen as the k -value which is considered optimal for such studies (Sareen et al., 2023; Troyanskaya et al., 2001).

Soft Imputing (SI)

Soft Imputing is a matrix completion technique that iteratively generates a lower-rank imputed matrix while minimizing the Frobenius norm difference from the original. It achieves so iteratively, by minimizing the following objective function:

$$\|X - M\|_F^2 + \lambda \|M\|_*$$

In Equation 3, X is the original matrix containing the observed values of Wind Speed and Flow Direction, M is the imputed matrix, $\|M\|_*$ is the nuclear norm of M , and λ is the regularization parameter. The above form is adopted from (Mazumder et al., n.d.).

In this paper, the value of λ is determined automatically, for each location, by the Python library ‘fancy impute 0.7.0’ (Feldman, n.d.).

Generative Adversarial Network (GAN)

GAN was first proposed in 2014 (Goodfellow et al., 2014). Since then, it has taken the AI community by storm. So much so that, the artwork “Portrait of Edmond de Belamy”, created using GAN was sold for \$ 432,000.00 (Jones, 2018), which is considered to be a milestone event in the history of evolution of Artificial Intelligence (Epstein et al., 2020).

GAN operates by pitting two neural networks against each other. The two competing neural networks are Generator and Discriminator.

Table 2: Summary of data considered in the study collected from NCEI

Station	Latitude (°)	Longitude (°)	Elevation relative to mean sea level (m)	Total number of observation	Percentage of missing data
Ahmedabad	23.0772	72.6346	57.6000	41305	21.4510
Indore	22.7218	75.8011	563.8800	28056	46.6464
Mangaluru	12.9613	74.8901	102.7100	39263	25.3342
Mumbai	19.0887	72.8679	11.2700	45643	13.2015

The generator network creates fake data, mimicking the actual data. Discriminator network tries to distinguish between the fake data and the actual data. When the fake and the actual data become indistinguishable from the Discriminator, then we say that the GAN model has been properly trained. The trained model is then deployed to create more synthetic data.

The generator tries to fool the Discriminator network, into believing that the fake data it produced, is in fact the real ones by minimising the loss function as shown in Equation 4. The discriminator network tries to identify the fake data by minimising the loss function as shown in Equation 5.

$$\text{Generator}_{\text{loss}} = \sum_{i=1}^n \log(D(G)(m_i)) \quad (4)$$

$$\text{Discriminator}_{\text{loss}} = \sum_{i=1}^n (\log(D(x_i)) + \log(1 - D(G(m_i)))) \quad (5)$$

In Equation 4 and Equation 5, n is the number of data points where the values of Wind Speed are missing. m_i is the noise input to the model, to prevent it from overfitting (Creswell & Bharath, 2019). x_i is the observed data of Wind Speed at the i^{th} data point. $G(m_i)$ is the fake Wind Speed data created by the Generator. $D(x_i)$ is a value between 0 and 1, denoting how well the discriminator thinks that x_i is indeed real. $D(G(m_i))$ is also a value between 0 and 1, which denotes how well the Discriminator thinks that $G(m_i)$ is also real.

Both Equation 4 and Equation 5 are minimized iteratively, using the Adam optimiser. Adam is often used in training GAN networks, as it is computationally efficient in handling large data (Kingma & Ba, 2014). Adam is a stochastic-gradient-based optimisation algorithm, which adjusts the weights of both Generator and Discriminator networks using the following equation:

$$w_{\text{new}} = w_{\text{old}} - l_{\text{rate}} \times m_t / \sqrt{V_t} + \varepsilon \quad (6)$$

In Equation 6, the new weight w_{new} , is calculated from the old weight w_{old} . New weights are calculated by taking a step along the gradient of the loss function, at the current step, m_t . The step size is controlled by learning rate, l_{rate} . To avoid the model, getting stuck in local minima, the velocity term v_t is adjusted (Shulman, 2023). A small number, ε , avoids division by 0 (Bhandari et al., 2021), which in turn avoids infinite learning rate.

GAN – Dual Annealing Algorithm

In this paper, a hybrid approach has been developed, which combines both GAN and Dual Annealing Algorithm (DA). Initially, GAN is employed to provide approximate values for the missing wind data.

These values are then further optimised to be a realistic representation of the wind speed. Optimization problems are those problems where objective function(s) must be satisfied, to arrive at an optimal solution. In addition to satisfying the objective function, if some constraints too are to be satisfied, then the problem becomes a Constrained Optimization Problem.

In this paper DA, from python library ‘scipy 1.11.2’ (Scipy, n.d.), has been adopted. DA is a stochastic optimization algorithm (Xiang et al., 1997). It is designed to achieve global minima (Silveira et al., 2022) by combining a simulated annealing algorithm with a local search algorithm. The objective functions considered in the study are:

$$\text{Minimize } |\text{Mean}_X - \text{Mean}_M| \quad (7)$$

$$\text{Minimize } |\text{Variance}_X - \text{Variance}_M| \quad (8)$$

$$\text{Minimize } |\text{Standard Deviation}_X - \text{Standard Deviation}_M| \quad (9)$$

While simultaneously solving the objective functions Equation 7 to Equation 9, the following constraints too are to be satisfied:

$$0 \leq \text{Wind Speed} \leq 90 \text{ m/s} \quad (10)$$

The constraint Equation 10 is in accordance with the data documentation (*FEDERAL CLIMATE COMPLEX*, n.d.).

Statistical Tools used to Assess the Performance of the Approaches

Mean

Mean of the observed and imputed data are calculated using the following formulae:

$$\text{Mean}_X = \frac{1}{N} \sum_{i=1}^N X \quad (11)$$

$$\text{Mean}_M = \frac{1}{N} \sum_{i=1}^N M \quad (12)$$

In Equation 11, N represents the total number of observations.

Variance

Variance of the observed and imputed data are calculated using the following formulae:

$$\text{Variance}_X = \frac{1}{N} \sum_{i=1}^N (X_i - \text{Mean}_X)^2 \quad (13)$$

$$\text{Variance}_M = \frac{1}{n} \sum_{i=1}^n (M_i - \text{Mean}_M)^2 \quad (14)$$

Standard Deviation

Standard deviation of the observed and imputed data is calculated using the following formulae:

$$\text{Standard Deviation}_X = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \text{Mean}_X)^2} \quad (15)$$

$$\text{Standard Deviation}_M = \sqrt{\frac{1}{n} \sum_{i=1}^n (M_i - \text{Mean}_M)^2} \quad (16)$$

Kurtosis

Kurtosis of the observed and imputed data are calculated using the following formulae:

$$\text{Kurtosis}_X = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \text{Mean}_X)^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \text{Mean}_X)^2 \right)^2} - 3 \quad (17)$$

$$\text{Kurtosis}_M = \frac{\frac{1}{n} \sum_{i=1}^n (m_i - \text{Mean}_M)^4}{\left(\frac{1}{n} \sum_{i=1}^n (m_i - \text{Mean}_M)^2 \right)^2} - 3 \quad (18)$$

Skewness

Skewness of the observed and imputed data are calculated using the following formulae:

$$\text{Skewness}_X = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \text{Mean}_X)^3}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \text{Mean}_X)^2 \right)^{\frac{3}{2}}} \quad (19)$$

$$\text{Skewness}_M = \frac{\frac{1}{n} \sum_{i=1}^n (m_i - \text{Mean}_M)^3}{\left(\frac{1}{n} \sum_{i=1}^n (m_i - \text{Mean}_M)^2 \right)^{\frac{3}{2}}} \quad (20)$$

R-Square

R-square is nothing but the proportion of variance in the independent variable, which can be predicted using the dependent variable (Chicco et al., 2021). However, in this case, since the matrix M has numerical values only in those places where the numbers are missing in the X matrix, it becomes difficult to calculate R-square directly for the two datasets.

Hence, the R-square value has been calculated based on the other statistical parameters, i.e., mean, variance, standard deviation, kurtosis, and skewness for the X and M matrices, and from that, the conclusion has been drawn.

Here the R-square value has been calculated using the following formula (Gugliani, 2020):

$$R^2 = 1 - \frac{\sum_{i=1}^5 (P_i - FF_i)^2}{\sum_{i=1}^5 (P_i - \overline{FF})^2} \quad (21)$$

In Equation 21, P_i is the cumulative distribution function (CDF) of the statistical values for the observed wind data. $\overline{FF}$ is the mean and F_i is the CDF of the statistical values for the imputed wind data.

Results and Discussion

Table 3 summarises the results of the statistical investigation of wind speed imputation for all four stations and methods considered in the study. In Table 3 and Figures 1 to 4, the original data is the observed data with missing values.

From Table 3 and Figures 1 to 4, we found that the *k-nn* imputation method is best at capturing the central tendency of the original data, as it has the same mean as the original data. However, the difference in the values of the other considered parameters indicates that the method fails miserably in capturing the variation and shape of the original distribution. This method works in the same manner everywhere, regardless of location.

For Ahmedabad, Mangaluru, and Mumbai, the *SI* imputation method captures neither the central tendency nor the variation and shape of the original distribution, as shown by the variation in the values of the considered parameters between the original and imputed distributions. However, for Indore, which has the highest percentage of missing data, the method could successfully capture the shape of the original distribution, as shown by the least variation in *kurtosis*, *skewness* between the original and imputed distributions. But, even there, it could not capture any

Table 3: Summary of statistical investigation for wind speed imputation

Station	Parameter	Mean	Variance	Standard Deviation	Kurtosis	Skewness	R-square
Ahmedabad	Wind Speed Original (m/s)	2.7993366421	1.4320324857	1.1966755975	0.7497612561	0.8096238881	
	Wind Speed k-nn (m/s)	2.7993366420	1.1248474246	1.0605882446	1.7737674358	0.9135092295	0.6312796478
	Wind Speed SI (m/s)	2.3061063041	2.0156738474	1.4197442894	-0.0067702069	0.5537532067	0.6911944466
	Wind Speed GAN (m/s)	2.8090154550	1.4456611547	1.2023565007	0.8428243729	0.8474296371	0.9986569206
Indore	Wind Speed GAN_DA_Optimized (m/s)	2.8061519512	1.4429484065	1.2012278745	0.8398841756	0.8458291123	0.9987093361
	Wind Speed Original (m/s)	3.2610635871	2.6275928808	1.6209851575	1.5809943388	1.1167642637	
	Wind Speed k-nn (m/s)	3.2610635871	1.4019158669	1.1840252814	5.5858151008	1.5289017256	0.0023143416
	Wind Speed SI (m/s)	1.9731273177	3.2992099689	1.8163727505	1.3700943713	1.2687663209	0.3870254620
Mangaluru	Wind Speed GAN (m/s)	3.2156469293	2.8334843884	1.6832956925	2.3495710233	1.3157454983	0.8728671216
	Wind Speed GAN_DA_Optimized (m/s)	3.1606517565	2.7396089168	1.6551764005	2.2177971154	1.2933317346	0.9046885604
	Wind Speed Original (m/s)	2.9184881441	1.8526172085	1.3611088158	0.9982711285	1.0497219491	
	Wind Speed k-nn (m/s)	2.9184881441	1.3832710746	1.1761254502	2.3548541367	1.2148243190	0.3651071939
Mumbai	Wind Speed SI (m/s)	2.3057830180	2.4896835863	1.5778731211	0.3887210409	0.8143187245	0.6497690648
	Wind Speed GAN (m/s)	2.9220264614	1.9360678466	1.3914265509	1.2699624298	1.1293358602	0.9872920145
	Wind Speed GAN_DA_Optimized (m/s)	2.9058294676	1.9131222298	1.3831566180	1.2238298334	1.1201302708	0.9915042926
	Wind Speed Original (m/s)	3.6114475385	2.4148602997	1.5539820783	0.3994041016	0.6747778451	
	Wind Speed k-nn (m/s)	3.6114475385	2.0960629202	1.4477786157	0.9164304830	0.7242769486	0.9536997750
	Wind Speed SI (m/s)	3.2006903109	3.2053909849	1.7903605740	-0.0813454156	0.3904206409	0.8839199300
	Wind Speed GAN (m/s)	3.6015155801	2.4454788016	1.5638026735	0.4443443484	0.7180648707	0.9999086853
	Wind Speed GAN_DA_Optimized (m/s)	3.5984673376	2.4418565568	1.5626440915	0.4397959078	0.7165182263	0.9999186451

Table 4: Instances of unrealistic data by GAN

Station	Date	Original	k-nn	SI	GAN	GAN_DA_Optimized
Indore	06-01-2018		3.261063587	0.5	-0.13269661	0.137220503
	08.00 PM					
Ahmedabad	05-02-2018		2.799336642	0.5	-0.088365108	0.052583573
	03.00 PM					

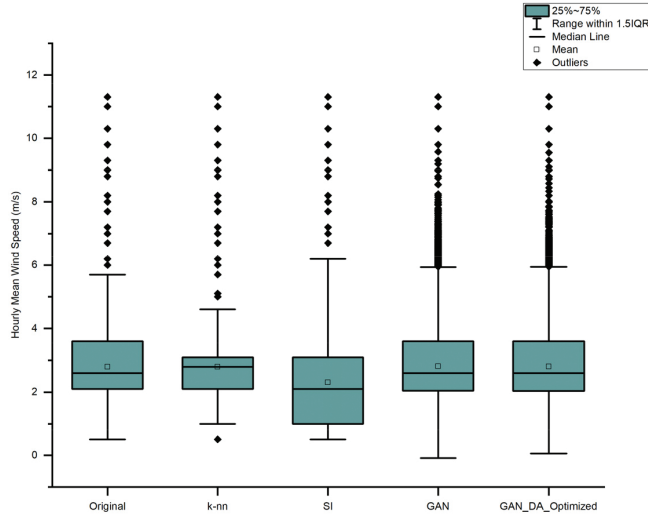


Figure 1: Box plot of wind speed distribution in Mangaluru.

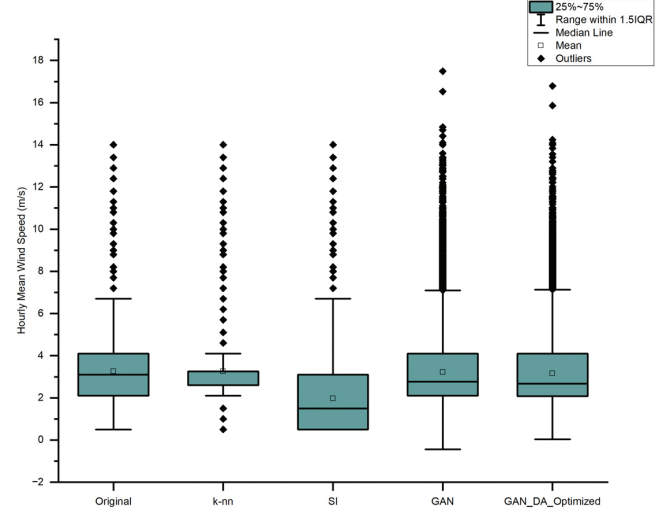


Figure 2: Box plot of wind speed distribution in Indore.

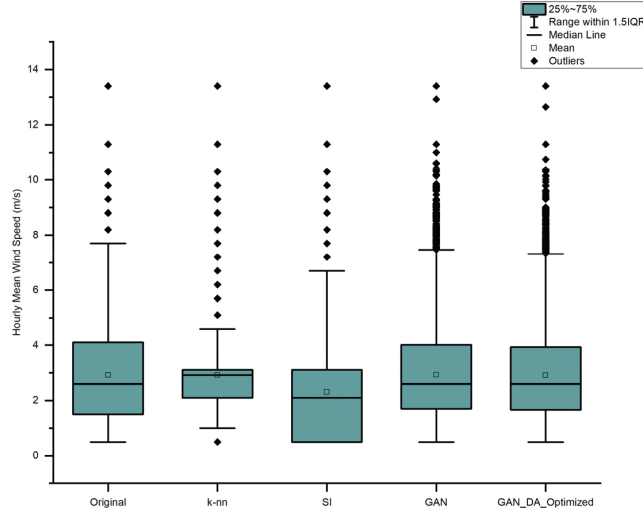


Figure 3: Box plot of wind speed distribution in Mangaluru.

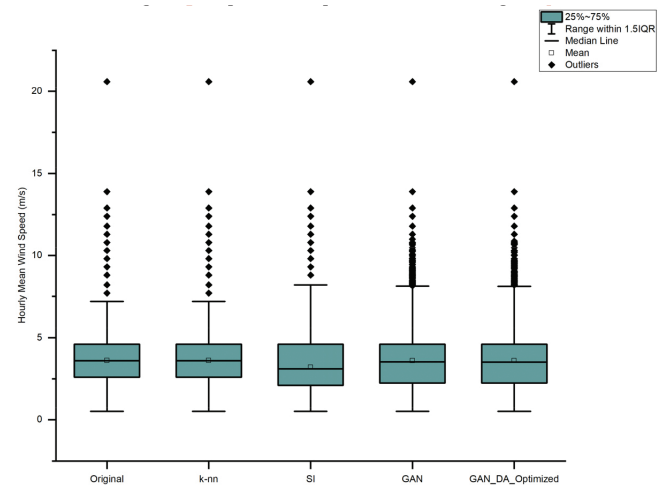


Figure 4: Box plot of wind speed distribution in Mumbai.

other attribute of the original distribution. Hence, we conclude that this method is the least suitable of all the considered methods to impute missing wind speed data for the considered stations.

Plain *GAN*, though not the best imputation method, could capture the central tendency, variation and shape of the original distribution to a fair degree of accuracy, which is indicated by high values of R-square, for all the places considered in the study. However, for Indore and Ahmedabad, it did generate unrealistic wind speed data (negative), as shown in Table 4, in certain instances. As such, this method is not suitable for imputing missing wind speed data.

To tackle this problem of *GAN*, we propose a hybrid approach. We take the output of plain *GAN* as input

for a constrained optimisation problem. The constraint in Equation 10 ensures that the generated values are all within the accepted range. Though the objective function of the optimisation problem is just to minimise the difference between *mean*, *variance*, and *standard deviation* values between the original and the imputed distribution, the method does successfully capture the shape of the original distribution, as shown by the least variation for *skewness* and *kurtosis*, in Table 3. From Table 4, we can confirm that in instances where plain *GAN* was producing unrealistic results, our method could predict values that satisfy all the conditions of the data.

Subsequently, our method scores the highest R-square value, irrespective of the location of the station, and hence can be considered the best performing method

for imputing missing wind speed data among all the considered methods in the four locations considered in the study.

Conclusions

Climate models are crucial tools for understanding and predicting climate change. Wind speed data are a key input for these models, but it can be missing due to various factors. This can introduce uncertainty into the models and lead to inaccurate estimations. Hence, in recent years, imputing missing wind data has gained traction.

This paper proposes a hybrid approach that uses both the GAN and dual annealing algorithms to impute missing wind speed data and counter unrealistic GAN results. We evaluated the performance of our approach against those of *k-nn*, *Soft Imputation*, and *plain GAN-based* methods on mean, variance, standard deviation, kurtosis, skewness, and R-square.

The hourly mean wind data for four Indian locations, viz. Ahmedabad, Indore, Mangaluru, and Mumbai, were obtained from the National Centers for Environmental Information. The study's observation period runs from January 1, 2017, at 12:00 AM through January 1, 2023 at 12:00 AM. The data are gathered at a height of 10 metres above ground level near airports to verify that the wind flow is not obstructed by local terrain or any other impediment.

Our results showed that our approach ranks number one based on the R-square value for all the considered stations. Our model consistently produces realistic results, unlike plain GAN. The hybrid approach is also able to capture the central tendency and spread of the data about the mean and shape of the original distribution.

Overall, our hybrid approach is a promising method for imputing missing wind speed data. It can produce realistic and accurate results, even when the percentage of missing data is high.

Future work could explore the application of our hybrid approach to impute missing wind speed data for other regions and time periods. Additionally, we could investigate other machine-learning methods to improve the performance of our approach.

References

- Afrifa-Yamoah, E., Mueller, U.A., Taylor, S.M. and Fisher, A.J., 2020. Missing data imputation of high-resolution temporal climate time series data. *Meteorological Applications*, **27**(1): e1873. <https://doi.org/10.1002/met.1873>
- Ammar, E. and Xydis, G., 2023. Wind speed forecasting using deep learning and preprocessing techniques. *International Journal of Green Energy*, **0**(0): 1-29. <https://doi.org/10.1080/15435075.2023.2228878>
- Ayiah-Mensah, F., Minkah, R., Asiedu, L. and Mettle, F.O., 2021. An enhanced method for tail index estimation under missingness. *Journal of Applied Mathematics*, **2021**: 1-13. <https://doi.org/10.1155/2021/3572555>
- Batista, G.E. and Monard, M.C., 2002. A study of K-nearest neighbour as an imputation method. Paper present at the Soft Computing Systems - Design, Management and Applications, HIS 2002, December 1-4, 2002, Santiago, Chile. **87**: 251-260.
- Battaglia, F., Cucina, D. and Rizzo, M., 2020. Detection and estimation of additive outliers in seasonal time series. *Computational Statistics*, **35**(3): 1393-1409. <https://doi.org/10.1007/s00180-019-00928-5>
- Ben, Y., Mei, Y., Chen, Y., Hu, T. and Zhu, D., 2020. Inter- and intra-annual wind speed variabilities in wide valley regions of the middle reaches of the Yarlung Tsangpo River, China. *Scientific Reports*, **10**(1): 12657. <https://doi.org/10.1038/s41598-020-69392-2>
- Bertsimas, D., Pawlowski, C. and Zhuo, Y.D., 2017. From predictive methods to missing data imputation: An optimization approach. *J. Mach. Learn. Res.*, **18**(1): 7133-7171.
- Bhandari, M., Parajuli, P., Chapagain, P. and Gaur, L., 2021. Evaluating performance of adam optimization by proposing energy index. In: Santosh, K., Hegadi, R. and Pal, U. (eds.). Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2021. Communications in Computer and Information Science, Springer, Cham, **1576**: 156-168. https://doi.org/10.1007/978-3-031-07005-1_15
- Boomgard-Zagrodnik, J.P. and Brown, D.J., 2022. Machine learning imputation of missing Mesonet temperature observations. *Computers and Electronics in Agriculture*, **192**: 106580. <https://doi.org/10.1016/j.compag.2021.106580>
- Brophy, E., Wang, Z., She, Q. and Ward, T., 2023. Generative adversarial networks in time series: A systematic literature review. *ACM Computing Surveys*, **55**(10): 199:1-199:31. <https://doi.org/10.1145/3559540>
- Burgette, L.F. and Reiter, J.P., 2010. Multiple imputation for missing data via sequential regression trees. *American Journal of Epidemiology*, **172**(9): 1070-1076. <https://doi.org/10.1093/aje/kwq260>
- Cai, J.-F., Candès, E.J. and Shen, Z., 2010. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, **20**(4): 1956-1982.
- Charakopoulos, A., Karakasidis, T. and Sarris, L., 2019. Pattern identification for wind power forecasting via complex network and recurrence plot time series analysis.

Afrifa-Yamoah, E., Mueller, U.A., Taylor, S.M. and Fisher, A.J., 2020. Missing data imputation of high-resolution

- Energy Policy*, **133**: 110934. <https://doi.org/10.1016/j.enpol.2019.110934>
- Cheng, C.S., Lopes, E., Fu, C. and Huang, Z., 2014. Possible impacts of climate change on wind gusts under downscaled future climate conditions: Updated for Canada. *Journal of Climate*, **27(3)**: 1255-1270. <https://doi.org/10.1175/JCLI-D-13-00020.1>
- Chicco, D., Warrens, M.J. and Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, **7**: e623. <https://doi.org/10.7717/peerj-cs.623>
- Creswell, A. and Bharath, A.A., 2019. Inverting the generator of a generative adversarial network. *IEEE Transactions on Neural Networks and Learning Systems*, **30(7)**: 1967-1974. <https://doi.org/10.1109/TNNLS.2018.2875194>
- Data Search | National Centers for Environmental Information (NCEI). (n.d.). Retrieved 25 July 2023, from <https://www.ncei.noaa.gov/access/search/data-search/global-hourly>
- Doddy Clarke, E., Sweeney, C., McDermott, F., Griffin, S., Correia, J.M., Nolan, P. and Cooke, L., 2022. Climate change impacts on wind energy generation in Ireland. *Wind Energy*, **25(2)**: 300-312. <https://doi.org/10.1002/we.2673>
- Donner, A., 1982. The relative effectiveness of procedures commonly used in multiple regression analysis for dealing with missing values. *The American Statistician*, **36(4)**: 378-381. <https://doi.org/10.1080/00031305.1982.10483055>
- Duvall, M.S., Jarvis, B.M. and Wan, Y., 2022. Impacts of climate change on estuarine stratification and implications for hypoxia within a shallow subtropical system. *Estuarine, Coastal and Shelf Science*, **279**: 108146. <https://doi.org/10.1016/j.ecss.2022.108146>
- Epstein, Z., Levine, S., Rand, D.G. and Rahwan, I., 2020. Who gets credit for AI-generated art? *Iscience*, **23(9)**: 101515.
- Farrugia, P.S. and Micallef, A., 2006. Comparative analysis of estimators for wind direction standard deviation. *Meteorological Applications*, **13(01)**: 29. <https://doi.org/10.1017/S1350482705001982>
- Faybishenko, B., Versteeg, R., Pastorello, G., Dwivedi, D., Varadharajan, C. and Agarwal, D., 2022. Challenging problems of quality assurance and quality control (QA/QC) of meteorological time series data. *Stochastic Environmental Research and Risk Assessment*, **36(4)**: 1049-1062. <https://doi.org/10.1007/s00477-021-02106-w>
- FEDERAL CLIMATE COMPLEX. (n.d.). Retrieved 26 July 2023, from <https://www.ncei.noaa.gov/data/global-hourly/doc/isd-format-document.pdf>
- Feldman, A.R., Sergey, n.d. *fancyimpute: Matrix completion and feature imputation algorithms* (0.7.0) [Python; OS Independent]. Retrieved 30 July 2023, from <https://github.com/iskandr/fancyimpute>
- Goodfellow, I., Bengio, Y. and Courville, A., 2016. *Deep learning*. MIT press.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, 27. Paper presented at the Proceedings of the 27th International Conference on Neural Information Processing Systems, **2**: 2672-2680. https://papers.nips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html
- Gugliani, G.K., 2020. Comparison of different multi-parameters probability density models for wind resources assessment. *Journal of Renewable and Sustainable Energy*, **12(6)**: 063303. <https://doi.org/10.1063/5.0024052>
- Hamer, R.M. and Simpson, P.M., 2009. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *The American Journal of Psychiatry*, **166(6)**: 639-641. <https://doi.org/10.1176/appi.ajp.2009.09040458>
- Han, S., Pool, J., Tran, J. and Dally, W., 2015. Learning both weights and connections for efficient neural network. *Advances in Neural Information Processing Systems*, **28**: 9.
- Jha, S.K., Wang, J. and Marina, N., 2022. A novel model-driven deterministic approach to wind power imputation. *Sustainable Computing: Informatics and Systems*, **36**: 100818. <https://doi.org/10.1016/j.suscom.2022.100818>
- Jing, B., Pei, Y., Qian, Z., Wang, A., Zhu, S. and An, J., 2022. Missing wind speed data reconstruction with improved context encoder network. *Energy Reports*, **8**: 3386-3394. <https://doi.org/10.1016/j.egy.2022.02.177>
- Jones, J., 2018. A portrait created by AI just sold for \$432,000: But is it really art? *The Guardian*. Available from <https://www.theguardian.com/artanddesign/shortcuts/2018/oct/26/call-that-art-can-a-computer-be-a-painter>
- Junger, W.L. and Ponce de Leon, A., 2015. Imputation of missing data in time series for air pollutants. *Atmospheric Environment*, **102**: 96-104. <https://doi.org/10.1016/j.atmosenv.2014.11.049>
- Kang, H., 2013. The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, **64(5)**: 402-406. <https://doi.org/10.4097/kjae.2013.64.5.402>
- Kayalvizhi, S. and D.M.V. Kumar, 2018. Stochastic Optimal Power Flow in Presence of Wind Generations Using Harmony Search Algorithm. In: 2018 20th National Power Systems Conference (NPSC), pp. 1-6. <https://doi.org/10.1109/NPSC.2018.8771822>
- Kingma, D.P. and Ba, J., 2014. Adam: A method for stochastic optimization. Preprint, Publisehd online on 22 December 2014. Available from arXiv Preprint arXiv:1412.6980.
- Kosana, V., Teeparthi, K. and Madasthu, S., 2022. A novel and hybrid framework based on generative adversarial network and temporal convolutional approach for wind speed prediction. *Sustainable Energy Technologies and Assessments*, **53**: 102467. <https://doi.org/10.1016/j.seta.2022.102467>

- Kumar, P.S., Muraleedharan, P.M., Prasad, T.G., Gauns, M., Ramaiah, N., de Souza, S.N., Sardesai, S. and Madhupratap, M., 2002. Why is the Bay of Bengal less productive during summer monsoon compared to the Arabian Sea? *Geophysical Research Letters*, **29**(24): 88-1-88-4. <https://doi.org/10.1029/2002GL016013>
- Lakku, N.K. and Behera, M.R., 2022. Skill and inter-model comparison of regional and global climate models in simulating wind speed over south Asian domain. *Climate*, **10**(6): 85.
- Liao, W., Bak-Jensen, B., Pillai, J.R., Yang, D. and Wang, Y., 2022. Data-driven missing data imputation for wind farms using context encoder. *Journal of Modern Power Systems and Clean Energy*, **10**(4): 964-976. <https://doi.org/10.35833/MPCE.2020.000894>
- Liu, H., Wang, Y. and Chen, W., 2020. Three-step imputation of missing values in condition monitoring datasets. *IET Generation, Transmission & Distribution*, **14**(16): 3288-3300. <https://doi.org/10.1049/iet-gtd.2019.1446>
- Liu, T., Wei, H. and Zhang, K., 2018. Wind power prediction with missing data using Gaussian process regression and multiple imputation. *Applied Soft Computing*, **71**: 905-916. <https://doi.org/10.1016/j.asoc.2018.07.027>
- Makwana, S. and Gandhi, N., n.d. Comparison between soil modulus based on standard penetration test and pressuremeter test- A case study of under ground Ahmedabad Metro. Available from <https://www.igs.org.in/storage/proceedings-uploads/TH11-32-170523110024.pdf>
- Malhotra, N.K., 1987. Analyzing marketing research data with incomplete information on the dependent variable. *Journal of Marketing Research*, **24**(1): 74-84. <https://doi.org/10.1177/002224378702400107>
- Mazumder, R., Hastie, T. and Tibshirani, R., 2010. Spectral regularization algorithms for learning large incomplete matrices. *The Journal of Machine Learning Research*, **11**: 2287-2322
- O'Reilly, C.H., Befort, D.J., Weisheimer, A., Woollings, T., Ballinger, A. and Hegerl, G., 2021. Projections of northern hemisphere extratropical climate underestimate internal variability and associated uncertainty. *Communications Earth & Environment*, **2**(1): 194. <https://doi.org/10.1038/s43247-021-00268-7>
- Outen, S. and Sobolowski, S., 2021. Extreme wind projections over Europe from the Euro-CORDEX regional climate models. *Weather and Climate Extremes*, **33**: 100363. <https://doi.org/10.1016/j.wace.2021.100363>
- Özen, C. and Deniz, A., 2022. Short-term wind speed forecast for Urla wind power plant: A hybrid approach that couples weather research and forecasting model, weather patterns and SCADA data with comprehensive data preprocessing. *Wind Engineering*, **46**(5): 1526-1549. <https://doi.org/10.1177/0309524X221088612>
- Pandey, D.M.M. (n.d.). Long-term Strategies and Programmes for Mechanization of Agriculture in Agro Climatic Zone-IX : Western Plateau and Hills region. pp. 187-197.
- Park, M.-S. and Baek, K., 2023. Quality management System for an IoT meteorological sensor network—application to smart seoul data of things (S-DoT). *Sensors*, **23**(5): 2384. <https://doi.org/10.3390/s23052384>
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V. and Gulin, A., 2018. CatBoost: Unbiased boosting with categorical features. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems, pp. 6639-6649.
- Qu, F., Liu, J., Ma, Y., Zang, D. and Fu, M., 2020. A novel wind turbine data imputation method with multiple optimizations based on GANs. *Mechanical Systems and Signal Processing*, **139**: 106610. <https://doi.org/10.1016/j.ymssp.2019.106610>
- Reichle, D.E., 2023. Climate and climate models. In: D.E. Reichle (Ed.), The Global Carbon Cycle and Climate Change, Second Edition, Chapter 13, Elsevier, pp. 389-452. <https://doi.org/10.1016/B978-0-443-18775-9.00010-3>
- Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature*, **323**(6088): 533-536. <https://doi.org/10.1038/323533a0>
- Sareen, K., Panigrahi, B.K., Shikhola, T. and Sharma, R., 2023. An imputation and decomposition algorithms based integrated approach with bidirectional LSTM neural network for wind speed prediction. *Energy*, **278**: 127799. <https://doi.org/10.1016/j.energy.2023.127799>
- Schewe, J., Gosling, S. N., Reyer, C., Zhao, F., Ciais, P., Elliott, J., Francois, L., Huber, V., Lotze, H.K., Seneviratne, S.I., van Vliet, M.T.H., Vautard, R., Wada, Y., Breuer, L., Büchner, M., Carozza, D. A., et al., 2019. State-of-the-art global models underestimate impacts from climate extremes. *Nature Communications*, **10**(1): 1005. <https://doi.org/10.1038/s41467-019-08745-6>
- Scipy: Fundamental algorithms for scientific computing in Python (1.11.2). (n.d.). [C, Python; MacOS, Microsoft :: Windows, POSIX, POSIX :: Linux, Unix]. Retrieved on 18 September 2023 from <https://scipy.org/>
- Shataee, S., Kalbi, S., Fallah, A. and Pelz, D., 2012. Forest attribute imputation using machine-learning methods and ASTER data: Comparison of k-NN, SVR and random forest regression algorithms. *International Journal of Remote Sensing*, **33**(19): 6254-6280. <https://doi.org/10.1080/01431161.2012.682661>
- Shulman, D., 2023. Optimization methods in deep learning: A comprehensive overview. Preprint. Published online 19 February 2023, arXiv:2302.09566; Version 1. arXiv. Available from <http://arxiv.org/abs/2302.09566>
- Silei, M., Bellavia, S., Superchi, F. and Bianchini, A., 2023. Recovering corrupted data in wind farm measurements: A matrix completion approach. *Energies*, **16**(4): 1674. <https://doi.org/10.3390/en16041674>
- Silveira, P., Teixeira, A.P. and Guedes Soares, C., 2022. A method to extract the quaternion ship domain parameters

- from AIS data. *Ocean Engineering*, **257**: 111568. <https://doi.org/10.1016/j.oceaneng.2022.111568>
- Stevens, B. and Bony, S., 2013. What are climate models missing? *Science*, **340(6136)**: 1053-1054. <https://doi.org/10.1126/science.1237554>
- Sun, B. and Tchetgen Tchetgen, E.J., 2018. On inverse probability weighting for nonmonotone missing at random data. *Journal of the American Statistical Association*, **113(521)**: 369-379. <https://doi.org/10.1080/01621459.2016.1256814>
- Tan, C.-H., Hou, J. and Chau, L.-P., 2015. Motion capture data recovery using skeleton constrained singular value thresholding. *The Visual Computer*, **31(11)**: 1521-1532. <https://doi.org/10.1007/s00371-014-1031-5>
- Tawn, R., Browell, J. and Dinwoodie, I., 2020. Missing data in wind farm time series: Properties and effect on forecasts. *Electric Power Systems Research*, **189**: 106640. <https://doi.org/10.1016/j.epsr.2020.106640>
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D. and Altman, R.B., 2001. Missing value estimation methods for DNA microarrays. *Bioinformatics*, **17(6)**: 520-525. <https://doi.org/10.1093/bioinformatics/17.6.520>
- Voosen, P. (29 July 2020). Missed wind patterns are throwing off climate forecasts of rain and storms: Climate models could improve by capturing hidden predictability. *Science: News*. Available from <https://doi.org/10.1126/science.abe0713>
- Wills, R.C.J., Dong, Y., Proistosecu, C., Armour, K.C. and Battisti, D.S., 2022. Systematic climate model biases in the large-scale patterns of recent sea-surface temperature and sea-level pressure change. *Geophysical Research Letters*, **49(17)**: e2022GL100011. <https://doi.org/10.1029/2022GL100011>
- Xi, D., Lin, N. and Gori, A., 2023. Increasing sequential tropical cyclone hazards along the US East and Gulf coasts. *Nature Climate Change*, **13(3)**: 258-265. <https://doi.org/10.1038/s41558-023-01595-7>
- Xiang, Y., Sun, D.Y., Fan, W. and Gong, X.G., 1997. Generalized simulated annealing algorithm and its application to the Thomson model. *Physics Letters A*, **233(3)**: 216-220. [https://doi.org/10.1016/S0375-9601\(97\)00474-X](https://doi.org/10.1016/S0375-9601(97)00474-X)
- Yao, Z., Wu, L., Huang, J. and Wang, X., 2023. Multivariate deep learning for reconstruction of spatial missing climate data. Paper presented at the 3rd International Conference on Artificial Intelligence, Automation, and High-Performance Computing (AIAHPC 2023), 12717: 474-482. <https://doi.org/10.1117/12.2684667>
- Yassen, A.H. and Abdul Kareem, A.K., 2023. Treatment missing data of daily and monthly air temperature in Iraq by using mean method. *Iraqi Journal of Science*, **64(3)**: 1566-1580. <https://doi.org/10.24996/ij.s.2023.64.3.44>
- Ye, X., Lu, Z., Qiao, Y., Min, Y. and O'Malley, M., 2016. Identification and correction of outliers in wind farm time series power data. *IEEE Transactions on Power Systems*, **31(6)**: 4197-4205. <https://doi.org/10.1109/TPWRS.2015.2512843>

Advertisement

Journal of Climate Change

[www.iospress.com/
journal-of-climate-change](http://www.iospress.com/journal-of-climate-change)



Aims and Scope

Climate change is reality which deals with the problem of climate variability and change and it deals with descriptions, causes, implications, interactions, impact and responses among other causes. The purpose of the journal is to provide a platform to exchange ideas among those working in different disciplines related to climate variations. The journal also plants to create an interdisciplinary forum for discussion of evidence of climate change, its causes, its natural resource impacts and its human impacts. The journal will also explore technological, policy, economy, strategic and social responses to climate change. It will be peer-reviewed, supported by rigorous processes of criterion-referenced article ranking and qualitative commentary, ensuring that only standard accepted quality work of the greatest substance and highest significance is published.

Editor-in-Chief

Prof. AL Ramanathan
School of Environmental Sciences
Jawaharlal Nehru University
New Delhi-10067, India
Tel: 91-11-26704314
Email: jcc@capital-publishing.com

Subscription Information 2024

ISSN 2395-7611
1 Volume, 4 issues (Volume 10)
Institutional subscription (online only):
US\$ 372 / €327
Individual subscription (online only):
US\$ 100 / €80

IOS Press serves the information needs of scientific and medical communities worldwide. IOS Press now publishes more than 100 international journals and approximately 75 book titles each year on subjects ranging from computer sciences and mathematics to medicine and the natural sciences.

IOS
Press

IOS Press
Nieuwe Hemweg 6B
1013 BG Amsterdam
The Netherlands
Tel.: + 31 20 688 3355
Fax: + 31 20 687 0019
Email: market@iospress.nl
URL: www.iospress.com

IOS Press c/o Accucoms US, Inc.
For North America Sales and Customer Service
West Point Commons
1816 West Point Pike
Suite 125
Lansdale, PA 19446, USA
Tel.: +1 215 393 5026
Fax: +1 215 660 5042
Email: iospress@accucoms.com