



Mapping Knowledge Structures in Deep Learning For Wind Energy Forecasting: A Technical and Bibliometric Review

Abdullah Ashoor Subih^{1,2}, Mohd Helmi Mansor^{*1}, Johnny Koh Siaw Paw³

¹ Institute of Power Engineering, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia. Emails: pe21189@student.uniten.edu.my, ORCID: 0009-0001-9394-7740

² Department of Electronic and Communications Engineering, College of Engineering and Petroleum, Hadhramout University, Mukalla, Hadhramout, Yemen

¹ Institute of Power Engineering, Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia. Emails: mhelmi@uniten.edu.my, ORCID: 0000-0002-3366-4976

³ Institute of Sustainable Energy, Universiti Tenaga Nasional, Putrajaya Campus, Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia. Emails: johnnykoh@uniten.edu.my, ORCID: 0000-0002-6219-7377

Abstract

Accurate wind-energy forecasting is essential for grid reliability and renewable-energy integration. This review aims to synthesize deep-learning methods for wind-speed and wind-power forecasting while mapping the field's knowledge structure. A technical review was integrated with bibliometric analysis of 324 Scopus-indexed publications from 2014–2024 using OpenRefine, BiblioMagika, VOSviewer, and Biblioshiny for data cleaning and co-authorship, co-citation, and keyword analyses. Results show that LSTM, Bi-LSTM, decomposition-assisted hybrids, and attention-based models dominate the literature; China leads in attributed publication output, while several lower-volume countries achieve higher citation rates per paper. Emerging themes include Transformers, federated learning, explainable artificial intelligence, and physics-informed forecasting. Persistent limitations involve data standardization, reproducibility, chronological evaluation, uncertainty calibration, and cross-site validation. The review concludes that progress depends on transparent benchmarks, robust probabilistic evaluation, and deployment-oriented reporting rather than model complexity alone.

Keywords: wind-energy forecasting; deep learning; wind-speed forecasting; wind-power forecasting; bibliometric analysis; knowledge mapping.

1. Introduction

The transition toward low-carbon power systems has made wind energy one of the fastest-expanding generation technologies worldwide (Mbugua & Hiraga, 2025). Rapid growth in installed wind capacity has created a corresponding need for accurate forecasts that can support unit commitment, reserve allocation, market participation, storage coordination, and network security. Nevertheless, wind generation remains difficult to predict because atmospheric dynamics are nonlinear, non-stationary, and sensitive to local terrain, turbine interactions, and forecast horizon. Consequently, forecasting errors can propagate into balancing costs, reserve requirements, curtailment decisions, and voltage or frequency management (Rajaperumal & Christopher Columbus, 2025).

Traditional physical and statistical forecasting methods remain valuable baselines. Physical models use numerical weather prediction and flow-field information, whereas statistical models infer relationships from historical observations. However, linear autoregressive formulations such as ARMA and ARIMA have limited capacity to represent multi-scale, nonlinear, and spatiotemporal dependencies. Their performance may also degrade under regime changes, abrupt wind ramps, missing observations, and distribution shifts between training and deployment periods (Rajaperumal & Christopher Columbus, 2025).

Deep learning (DL) addresses these limitations through multilayer networks that learn hierarchical representations directly from raw data. Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and attention-based Transformer architectures have repeatedly outperformed conventional baselines in wind-forecasting studies (Wu et al., 2022; Faruque et al., 2024), and comparable gains have been reported for adjacent renewable resources such as photovoltaic generation (Verdone et al., 2025).

Several bibliometric studies have examined neighboring domains. Zhang et al. (2022) mapped artificial intelligence in renewable-energy systems; Ye et al. (2020) analyzed offshore wind-power research; Lagos et al. (2022) documented the prevalence of hybrid models in wind-speed and wind-power forecasting; and Zaidi (2024) examined deep-learning applications in energy production. These studies are bibliometric rather than technical in method, and their datasets do not extend beyond 2022, leaving newer federated and Transformer-based wind-forecasting models outside their coverage.

This review addresses the gap through two complementary components: (1) a technical synthesis of DL architectures, decomposition-assisted hybrids, uncertainty-aware models, and metaheuristic hyperparameter optimization; and (2) a bibliometric analysis of 324 Scopus-indexed publications from 2014–2024 covering publication growth, influential contributors, source venues, citation impact, keyword structure, and thematic evolution. The bibliometric corpus is fixed to 2014–2024; selected post-2024 studies are used only to update the technical discussion and are not included in the reported bibliometric indicators. The study is guided by six research questions concerning the current research landscape (RQ1), publication trends (RQ2), leading authors, institutions and countries (RQ3), source venues and high-impact documents (RQ4), thematic clusters (RQ5) and keyword co-occurrence evolution (RQ6). The remainder of the paper is organized as follows: Section 2 presents the theoretical background; Section 3 describes the materials and methods; Section 4 reports and discusses results; and Section 5 presents the conclusion.

2. Theoretical Background

2.1 Forecasting Horizons and Problem Formulation

Wind energy forecasting problems are classified by temporal horizon, each with distinct operational implications (Alkhayat & Mehmood, 2021): very-short-term (seconds to 30 min) for real-time turbine pitch control and electricity market settlements; short-term (30 min to 6 h) for unit commitment, economic dispatch, and reserve scheduling; medium-term (6 h to 1 week) for maintenance planning, energy storage management, and day-ahead trading; and long-term (beyond 1 week) for capacity planning and grid expansion decisions (Mbugua & Hiraga, 2025; Verdone et al., 2025). Physically, wind power output P is related to wind speed v through the aerodynamic power equation:

$$P = \frac{1}{2} \rho A C_p v^3 \quad (1)$$

where ρ denotes air density (kg/m^3), A is the turbine rotor swept area (m^2), and C_p is the dimensionless power coefficient (Betz limit: $C_p \leq 0.593$). The cubic relationship between power output and wind speed amplifies forecasting errors nonlinearly. Wake effects, velocity deficits downstream of turbines, further complicate power prediction at the wind farm scale, as intra-farm interactions generate complex spatial dependencies that cannot be captured by single-turbine models (Alkhayat & Mehmood, 2021). Forecasting tasks are further categorized as deterministic (single-point predictions with implicit uncertainty) or probabilistic (confidence intervals, quantile forecasts, and scenario generation), with the latter increasingly favored for risk-aware grid operations and energy market participation (Wang et al., 2020a; Niu et al., 2022; Afrasiabi et al., 2021).

2.2 Deep-Learning Architectures

The DL landscape for wind forecasting encompasses a diverse set of architectures with distinct inductive biases, each suited to different facets of the forecasting problem. Convolutional Neural Networks (CNNs) apply learned filter banks to extract local spatial and frequency-domain features from multi-channel meteorological fields and SCADA inputs, making them particularly effective for capturing turbine spatial correlation patterns across wind farms (Harbola & Coors, 2019; Yu et al., 2019; Zhu et al., 2018). Standard Recurrent Neural Networks (RNNs) model sequential temporal dependencies but suffer from the vanishing gradient problem when learning over long time horizons.

Long Short-Term Memory (LSTM) networks address this limitation through three learnable gates. The forget gate determines which information is removed from the cell state (Wu et al., 2022):

$$f_t = \sigma(W_f [h_{t-1}, x_t] + b_f) \quad (2)$$

The input gate determines which new information is written to the cell state:

$$i_t = \sigma(W_i [h_{t-1}, x_t] + b_i) \quad (3)$$

The output gate controls which information is exposed from the cell state:

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \quad (4)$$

Together, these gates support selective retention of long-range dependencies and capture multiday periodicity and seasonal structure (Mujeeb et al., 2019; Liu et al., 2018c). Bidirectional LSTM (Bi-LSTM) processes sequences in both temporal directions, enriching hidden states with past and future context and supporting multistep forecasting (Neshat et al., 2021; Xiong et al., 2023). Gated Recurrent Units simplify the mechanism to two gates and can achieve comparable accuracy on many wind-forecasting benchmarks with fewer parameters and faster convergence (Li et al., 2020).

Hybrid CNN-LSTM stacks exploit complementary strengths, with convolutional layers extracting local features from multivariate input windows and recurrent layers modeling the resulting feature sequences; this combination consistently outperforms single-architecture models across wind farm datasets (Wu et al., 2016; Wang et al., 2019). Attention mechanisms learn context-dependent weights across input time steps, allowing the model to concentrate on the most informative history for each horizon (Xiong et al., 2022; Ma & Mei, 2022).

The Transformer, built entirely on multi-head self-attention without recurrence, scales efficiently to long sequences and has reported state-of-the-art ultra-short-term performance (Bommidi et al., 2023). Autoencoder-based and ensemble formulations, which combine multiple predictors through stacking, boosting, or Bayesian model averaging, improve robustness to non-stationarity (Hong & Rioflorido, 2019; Ibrahim et al., 2021; Wang et al., 2021b; Aly, 2020). Spatiotemporal designs that jointly encode geographic and temporal structure extend these ideas to multi-site settings (Hong & Satriani, 2020; Liu et al., 2020c). Federated architectures represent the most recent addition (Li et al., 2023). Table 1 summarizes the principal architecture families.

Table 1. Comparative Summary of Major Deep-Learning Architectures for Wind-Energy Forecasting

Representative Study	Architecture	Key Property
Liu et al. (2018c)	LSTM	Long-range temporal memory via gating
Neshat et al. (2021)	Bi-LSTM	Bidirectional sequence processing
Li et al. (2020)	GRU	Efficient 2-gate recurrent unit
Wu et al. (2016)	CNN-LSTM	Spatial-feature + temporal hybrid
Xiong et al. (2023)	Attention-LSTM	Dynamic time-step weighting
Bommidi et al. (2023)	Transformer	Pure self-attention; no recurrence
Hong and Rioflorido (2019)	Ensemble DL	Multi-model stacking/boosting
Li et al. (2023)	Federated DL	Privacy-preserving multi-site

2.3 Preprocessing and Decomposition Methods

Wind speed series superimpose long-term seasonal trends, diurnal periodicity, and stochastic turbulence. Decomposing these components before training lets each sub-model specialize on a more stationary sub-signal, and has become standard practice (Alkhayat & Mehmood, 2021). Wavelet-based methods, including the Discrete Wavelet Transform, Wavelet Packet Decomposition and Maximal Overlap DWT, split the signal into orthogonal frequency sub-bands at multiple resolutions (Yu et al., 2018; Zhang et al., 2022c). Empirical Mode Decomposition and its ensemble extensions (EEMD, CEEMDAN, ICEEMDAN) adaptively derive Intrinsic Mode Functions without a predefined basis; ICEEMDAN specifically mitigates the mode mixing and frequency aliasing of earlier variants (Bommidi et al., 2023; Kosana et al., 2022).

Variational Mode Decomposition (VMD), posed as a constrained optimization that minimizes the summed bandwidth of the extracted modes, yields components with well-defined center frequencies and greater noise resistance than EMD-family methods (Wang et al., 2018; Yang et al., 2019). VMD-LSTM, VMD-BiLSTM and VMD-GRU hybrids report strong benchmark performance across diverse datasets (Wang & Li, 2018; Ma et al., 2020; Krishna Rayi et al., 2022; Tian, 2021). Singular Spectrum Analysis decomposes the trajectory matrix by singular value decomposition, separating trend, oscillatory, and noise components, and SSA-LSTM and SSA-GRU hybrids achieve competitive multistep accuracy (Liu et al., 2019; Yan et al., 2020).

Secondary or two-layer decomposition schemes, which decompose the residual of a first-stage decomposition, have been reported to improve accuracy further at the cost of added complexity (Xiang et al., 2020; Zhang et al., 2021; Yin et al., 2019; Liu et al., 2021). Related variants combine wavelet denoising with recurrent layers (Liu et al., 2020b), apply time-varying filter EMD with evolutionary CNN-BiLSTM stacks (Zhang et al., 2022a), or embed denoising within an ensemble predictor (Peng et al., 2020; Hossain et al., 2021). Feature selection and data augmentation are complementary preprocessing routes: multi-objective feature selection reduces input redundancy in multivariate settings (Lv & Wang, 2023), while time-series augmentation improves generalization when historical records are short (Bandara et al., 2021; Wang et al., 2021a).

2.4 Hyperparameter Optimization

DL forecasting accuracy is highly sensitive to hyperparameter configuration, including layer count, units per layer, learning rate, dropout rate, batch size, sequence length, and optimizer choice. The search space is high-dimensional and non-convex, which makes manual tuning impractical for production systems, and population-based metaheuristics have consequently been adopted as black-box optimizers (Gu et al., 2023). The Whale Optimization Algorithm, which mimics the encircling and bubble-net foraging of humpback whales, has been applied to LSTM and Bi-LSTM configurations (Huangfu et al., 2022; Samadianfard et al., 2020; Ding et al., 2022).

Particle Swarm Optimization, the Grey Wolf Optimizer, the Salp Swarm Algorithm, and Genetic Algorithms are also widely used. Coati Optimization Algorithm-tuned CNN-LSTM frameworks have demonstrated competitive performance, while LSTM and Bi-LSTM models have been optimized using WOA, PSO, and GA for wind-forecasting tasks (Abou Houran et al., 2023; Subih et al., 2026).

Crisscross and bi-attention formulations extend this to joint architecture-parameter search (Meng et al., 2022). Error-correction strategies, which fit a secondary model to the residuals of the primary predictor, reduce systematic bias in multistep forecasts, particularly during ramp events (Ma et al., 2020), and quantile regression extensions generate prediction intervals without distributional assumptions (Niu et al., 2022; Peng et al., 2021).

2.5 Evaluation Protocols and Reproducibility

Meaningful comparison requires evaluation protocols that preserve temporal order and separate point accuracy from uncertainty quality. Deterministic studies should report complementary metrics such as MAE, RMSE or normalized RMSE, and a scale-sensitive measure where appropriate. Probabilistic studies should additionally report calibration and sharpness through pinball loss, continuous ranked probability score, prediction interval coverage probability and interval width. Random train-test partitioning is inappropriate for time series; chronological holdout or rolling walk-forward validation gives a more credible estimate of operational performance (Peng et al., 2021).

Reproducibility further requires disclosure of data sources, sampling resolution, forecast horizon, preprocessing, missing-data handling, hyperparameter search bounds, random seeds, stopping criteria, baseline definitions and computational cost. For deployment-oriented work, inference latency, throughput, memory footprint, model size, recalibration frequency and robustness to distribution shift matter as much as accuracy. These elements determine whether a statistically accurate model is usable in real-time grid operation, multi-site transfer or embedded turbine-level implementation (Lin et al., 2020; Wu et al., 2022). Reporting of these items in the reviewed corpus is uneven, which is itself a finding of this review.

2.6 Prior Bibliometric Studies

Table 2 compares representative bibliometric studies of wind-energy forecasting and related artificial-intelligence applications. The present review differs in four respects: it focuses specifically on deep-learning-based wind-speed and wind-power forecasting; integrates technical synthesis with bibliometric analysis; analyzes a Scopus corpus covering 2014–2024; and combines BiblioMagika, OpenRefine, VOSviewer, and Biblioshiny for data harmonization and network visualization. These features provide complementary technical and bibliometric perspectives without implying direct superiority over prior single-tool studies.

Table 2. Comparison of Prior Bibliometric Studies Related to Wind Forecasting and AI-Enabled Renewable-Energy Research

Author/Year	Domain & Query	Database; Scope	Key Findings
Ye et al. (2020)	Offshore wind power	WoS; 1979–2018; n=1,562	UK leads; grid-connection and environmental impact are key themes.

Author/Year	Domain & Query	Database; Scope	Key Findings
Zhang et al. (2022)	AI in renewable energy	WoS; 2006–2022; n=469	China dominates; hybrid forecasting and system stability are core clusters.
Lagos et al. (2022)	Wind speed/power forecasting	VOSviewer; to 2021	Hybrid models dominate; short-term horizons are the primary focus.
Zaidi (2024)	DL in energy production	Scopus; 2004–2022; n=152	Grid forecasts, generation management, and energy optimization are identified.
Present study	DL for wind speed/power forecasting	Scopus; 2014–2024; n=324	Technical synthesis + bibliometric mapping; hybrid DL architectures dominate.

3. Materials and Methods

3.1 Database Selection and Search Query

A focused bibliometric investigation was conducted using Scopus, selected for its broad peer-reviewed coverage and structured metadata covering citations, authors, affiliations, sources and keywords. The query was restricted to the TITLE field to prioritize topical precision and retrieve publications whose primary subject is DL-based wind speed or wind power forecasting:

TITLE (((wind AND speed AND prediction) OR (wind AND speed AND forecasting) OR (wind AND power AND forecasting) OR (wind AND power AND prediction)) AND ((deep AND learning) OR (hybrid AND deep AND learning)))

Records were restricted to the 2014–2024 window, screened for title-level relevance and manually checked for alignment with wind speed or wind power forecasting using DL. Off-topic records were removed before analysis, yielding a final corpus of 324 documents. Selected 2025–2026 publications were consulted only for the narrative technical update in Section 2 and were excluded from all bibliometric counts and network maps. The title-only strategy reduces false positives but may exclude studies in which DL wind forecasting is a secondary contribution; results should therefore be read as a high-precision mapping of the field's core literature rather than an exhaustive retrieval. The workflow is summarized in Figure 1.

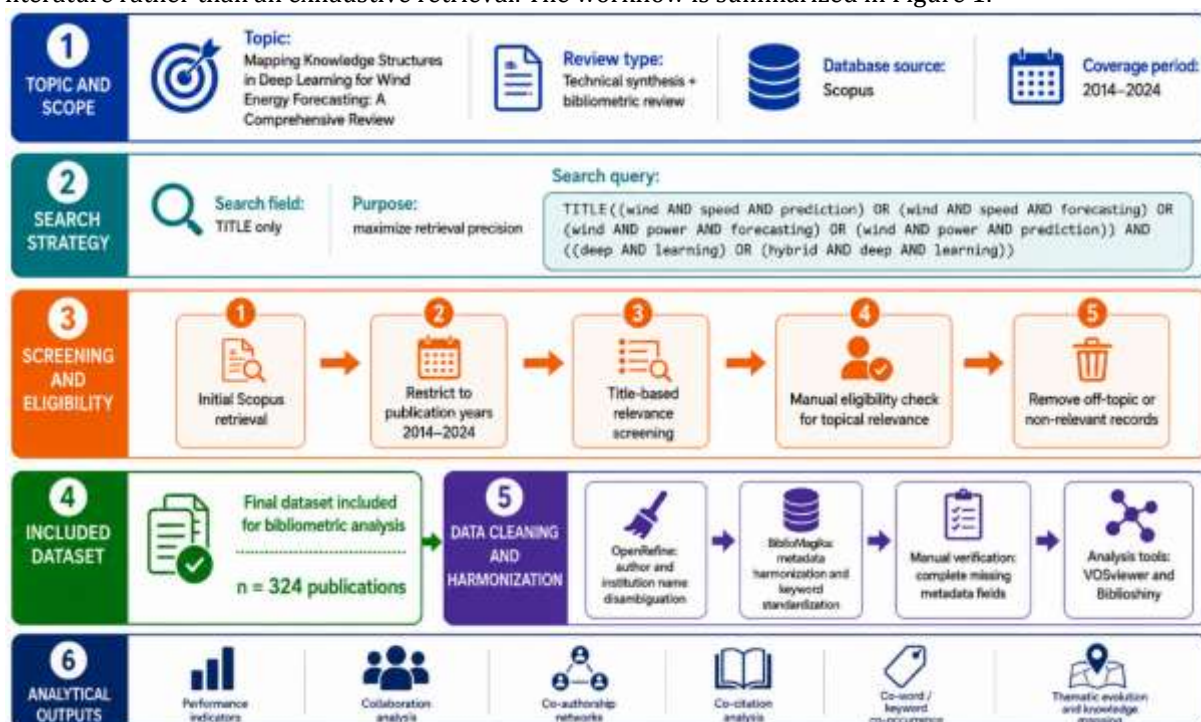


Figure 1. Search Strategy and Bibliometric Analysis Workflow

3.2 Data Cleaning and Bibliometric Indicators

Data cleaning and harmonization are prerequisites for reliable bibliometric inference (Donthu et al., 2021; Lim et al., 2024). OpenRefine (version 3.7) was used for cluster-based string matching to disambiguate author-name variants and inconsistent institutional names. BiblioMagika was then used for metadata harmonization, controlled-vocabulary keyword standardization, and generation of standardized bibliometric reports. Missing metadata fields were resolved through manual verification against the original Scopus records.

Indicators were organized into three groups. Performance indicators comprised total publications (TP), total citations (TC), citations per paper (C/P), the h-index, the g-index, and the m-index. The h-index is the largest value h for which h publications have each received at least h citations; the g-index is the largest value g for which the top g publications have collectively received at least g^2 citations; and the m-index is the h-index divided by the number of active publication years. Collaboration indicators comprised authors per paper (A/P), citations per author, and institutional and country-level contribution counts. Impact indicators comprised the number of cited publications, citations per cited paper (C/CP), and the annual citation rate (C/Y). Co-authorship, co-citation, and co-word networks were produced using VOSviewer (version 1.6.20) and Biblioshiny (Bibliometrix R package, version 4.1.4). Country-level figures use full counting; therefore, internationally co-authored papers are credited to every participating country, and attributed totals exceed the unique corpus size.

4. Results and Discussion

4.1 Current Research Landscape and Corpus Profile

The corpus is dominated by peer-reviewed journal articles (72.53%), with conference papers accounting for a further 25.31% (Table 3). Only one review is present, indicating limited synthesis within the domain and supporting the motivation for the present study. The near-exclusive use of English (93.83%) reflects Scopus indexing practice rather than the global distribution of research activity and is revisited as a limitation in Section 4.8.

Table 3. Corpus profile by document type, language and source type (n = 324).

Category	Class	TP	%
Document type	Article	235	72.53
	Conference paper	82	25.31
	Book chapter	3	0.93
	Erratum	2	0.62
	Review	1	0.31
	Retracted	1	0.31
Language	English	304	93.83
	Chinese	18	5.56
	Korean	1	0.31
	Turkish	1	0.31
Source type	Journal	241	74.38
	Conference proceeding	76	23.46
	Book series	6	1.85
	Book	1	0.31

TP = total publications.

Table 4 reports aggregate citation indicators. The corpus contains 324 documents and 10,174 citations, corresponding to 31.40 citations per paper and 39.90 citations per cited paper. Across the eleven-year

window, the average annual citation rate is 924.91 citations per year. The h-index of 56 and g-index of 94 indicate a substantial body of influential work, but these are corpus-level impact measures and are not direct evidence of methodological quality. The average of 4.54 authors per document reflects the collaborative character of the field.

Table 4. Aggregate citation indicators for the study corpus (2014-2024).

Indicator	Value
Total documents	324
Citable years	11
Contributing authors	1,471
Cited documents	255
Aggregate citations	10,174
Citations per paper (C/P)	31.40
Citations per cited paper (C/CP)	39.90
Annual citation rate (C/Y)	924.91
Authors per document (A/P)	4.54
h-index	56
g-index	94
m-index (h/n)	5.09

h-index; g-index; m-index = h/n .

4.2 Publication and Citation Trends

Table 5 and Figure 2 show growth from a single publication in 2014 to a peak of 76 in 2023, with contributing authors rising from 3 to 337 over the same interval. The cumulative h-index reaches 56. Average citations per paper fall from 148.67 in 2016 to 2.85 in 2024; this reflects the citation accumulation lag for recent work rather than a decline in quality, and any interpretation of the recent years must account for that truncation.

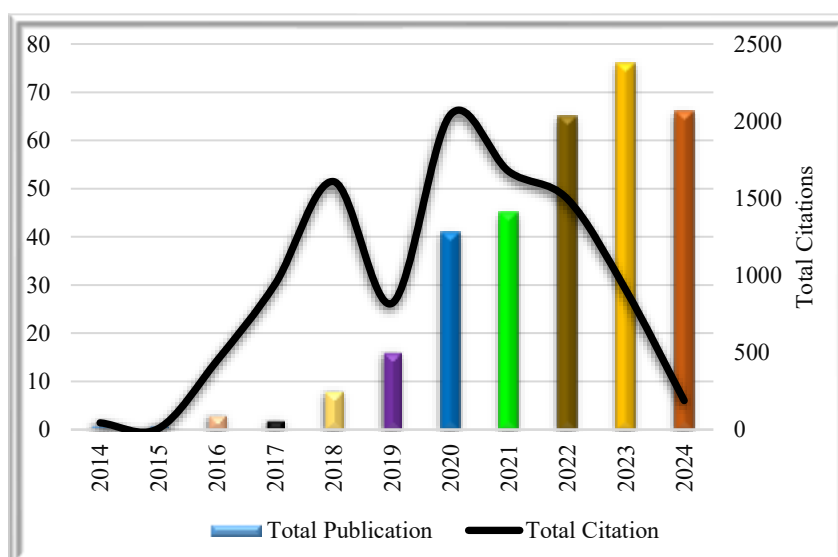


Figure 2. Annual Publication Output and Citation Accumulation, 2014-2024

Table 5. Annual publication and citation metrics (2014-2024).

Year	TP	NCA	TC	C/P	h	g	m
2014	1	3	44	44.00	1	1	0.09

Year	TP	NCA	TC	C/P	h	g	m
2015	1	5	7	7.00	1	1	0.10
2016	3	11	446	148.67	3	3	0.33
2017	2	10	945	472.50	2	2	0.25
2018	8	33	1,608	201.00	7	8	1.00
2019	16	73	819	51.19	12	16	2.00
2020	41	178	2,042	49.80	24	41	4.80
2021	45	212	1,673	37.18	25	40	6.25
2022	65	299	1,496	23.02	23	37	7.67
2023	76	337	906	11.92	16	28	8.00
2024	66	310	188	2.85	7	12	7.00
Total	324	1471	10,174	31.40	56	94	5.09

TP = total publications; NCA = contributing authors; TC = total citations.

4.3 Leading Authors and Institutions

Table 6 lists authors with at least eleven publications. Yin H. is the most prolific contributor with 24 publications, while Liu H. records the highest citation rate per paper in this cohort ($C/P = 92.26$). The table combines productivity with citation impact and should not be read as a ranking of research quality, since citation counts vary with publication age, subfield, and collaboration pattern.

Table 6. Leading authors with at least eleven publications.

Author	Country	TP	TC	C/P	h
Yin, H.	CN	24	1,296	54.00	12
Lai, Y.-H.	CN	19	732	38.53	11
Liu, H.	CN	19	1,753	92.26	14
Chen, C.	CN	17	197	11.59	7
Wang, Y.	IN	16	257	16.06	8
Wang, B.	CN	16	224	14.00	9
Zhang, C.	CN	15	477	31.80	8
Shao, H.	US	14	778	55.57	7
Li, Y.	CN	12	145	12.08	7
Peng, X.	CN	12	238	19.83	8
Zhang, H.	KR	12	438	36.50	6
Chen, H.	CN	11	624	56.73	7

CN = China; IN = India; US = United States; KR = South Korea.

Table 7 presents institutional productivity for organizations with at least ten publications. Huazhong University of Science and Technology leads in both volume ($TP = 75$) and total citations ($TC = 3,452$). Among the listed institutions, Kunming University of Science and Technology records the highest citations per paper ($C/P = 77.75$), followed by Galgotias College of Engineering and Technology ($C/P = 70.06$) and North China Electric Power University ($C/P = 60.24$). Volume and average influence are therefore distinct dimensions, and the appearance of Indian, British and Australian institutions alongside Chinese ones indicates a broadening institutional base.

Table 7. Leading institutions with at least ten publications.

Institution	Country	TP	TC	C/P	h
Huazhong U. Science and Technology	CN	75	3,452	46.03	25
Guangdong U. of Technology	CN	48	1,801	37.52	20
North China Electric Power U.	CN	42	2,530	60.24	23
Central South University	CN	32	1,310	40.94	18
Harbin Institute of Technology	CN	27	1,428	52.89	10
Dongbei U. Finance and Economics	CN	24	718	29.92	10
China Electric Power Research Inst.	CN	24	591	24.63	11
Huaiyin Institute of Technology	CN	17	446	26.24	8
Galgotias College Eng. and Tech.	IN	16	1,121	70.06	12
University of Glasgow	UK	13	195	15.00	4
Amrita University	IN	13	195	15.00	8
Kunming U. Science and Technology	CN	12	933	77.75	6

CN = China; IN = India; UK = United Kingdom.

4.4 Country-Level Contributions

Table 8 and Figure 3 present the geographical distribution under full counting. China records the highest attributed publication count (308 country-paper pairings), followed by India (104) and the United States (36). Because internationally co-authored papers are credited to every participating country, attributed totals are non-exclusive and exceed the unique corpus size. China leads in volume, whereas South Korea (C/P = 52.82), Hong Kong (48.60), Ethiopia (47.18) and Turkey (43.50) show higher citation rates per paper than China (32.63). These figures describe the distribution of indexed output and citation impact and should not be read as measures of national research quality

Table 8. Country-level bibliometric contributions (top twelve, full counting).

Country	TP	TC	C/P	h	g
China	308	10,051	32.63	55	100
India	104	3,173	30.51	29	56
United States	36	1,308	36.33	11	36
Australia	31	892	28.77	16	29
United Kingdom	29	679	23.41	9	26
Taiwan	21	412	19.62	10	20
Iran	20	260	13.00	7	16
Pakistan	18	330	18.33	7	18
South Korea	17	898	52.82	9	17
Turkey	16	696	43.50	7	16
Ethiopia	11	519	47.18	7	11
Hong Kong	10	486	48.60	6	10

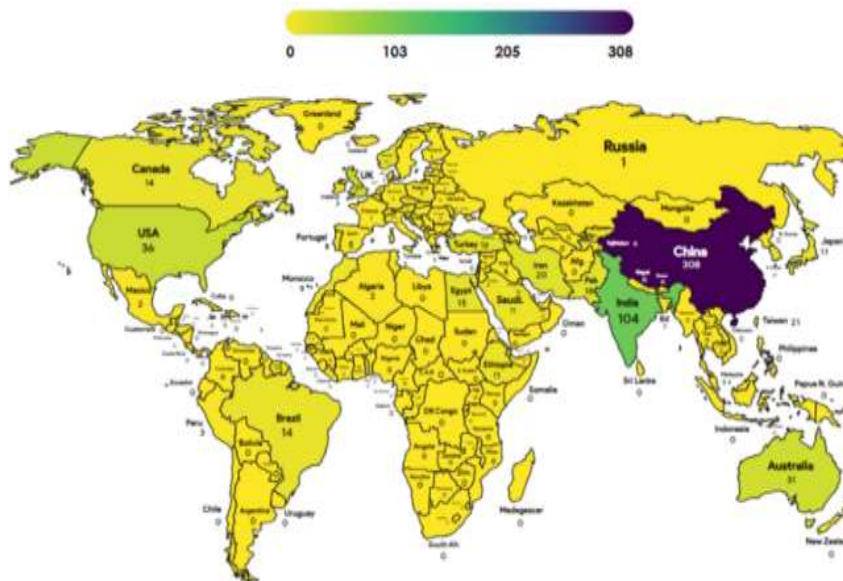


Figure 3. Worldwide scientific production in DL-based wind energy forecasting (Scopus, 2014-2024).

4.5 Source Titles and Highly Cited Documents

Table 9 lists leading source titles with at least five publications. Applied Energy has the highest publication count (TP = 38) and total citations (TC = 4,218), followed closely by Energy Conversion and Management (TP = 35; TC = 4,105), which records the highest citations per paper among these venues (C/P = 117.28). The concentration of output in applied energy and power systems journals confirms that the field is anchored in energy engineering rather than in core machine learning venues, with implications for how methodological rigor is assessed.

Table 9. Leading source titles with at least five publications.

Journal	TP	TC	C/P	h
Applied Energy	38	4,218	110.99	25
Energy Conversion and Management	35	4,105	117.28	24
Renewable Energy	29	2,241	77.28	18
Energy	27	1,847	68.41	16
Journal of Cleaner Production	18	1,021	56.72	14
IEEE Trans. Sustainable Energy	12	623	51.92	10
Energies	11	441	40.09	9
Electric Power Systems Research	10	387	38.70	8
Applied Soft Computing	8	498	62.25	7
Int. J. Electrical Power and Energy Syst.	7	302	43.14	6
IEEE Access	6	214	35.67	5
Energy Reports	5	98	19.60	4

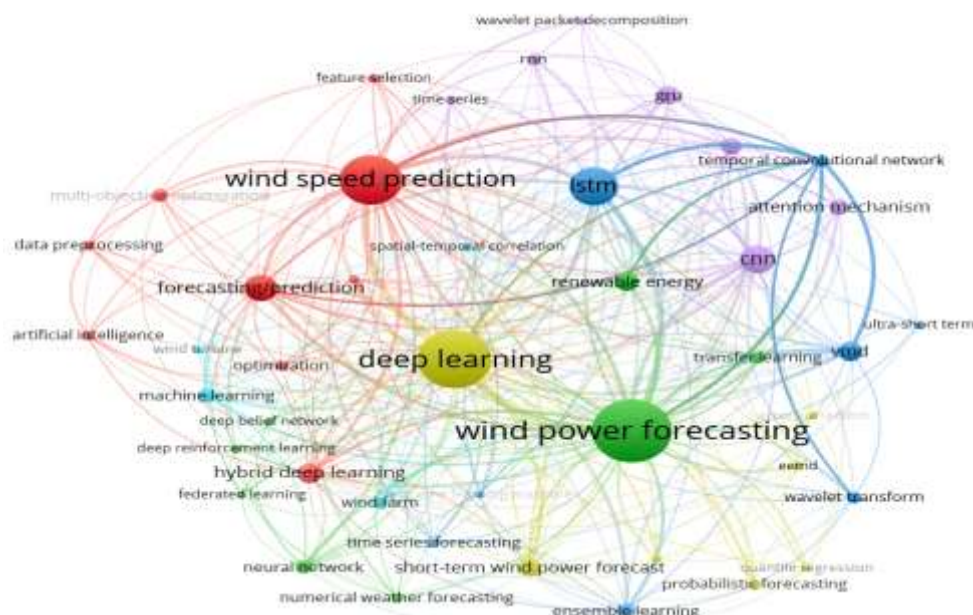
Table 10 lists the twelve most cited documents. Wang et al. (2017) lead with 627 citations, highlighting the influence of ensemble probabilistic forecasting. Three studies by Liu et al. (2018a, 2018b, 2018c) received 273, 405, and 405 citations, respectively. Hu et al. (2016) introduced transfer learning for short-term wind-speed prediction (377 citations), and Qureshi et al. (2017) applied meta-regression and transfer learning (318 citations). The remaining highly cited documents were authored by Chen et al. (2018), Hong and Rioflorido (2019), Jiang et al. (2021), Neshat et al. (2021), Wang et al. (2020b), and Li et al. (2023).

Table 10. Twelve Most Cited Documents in the Corpus

Reference	Year	Source	TC	C/Year
Wang et al. (2017)	2017	Applied Energy	627	78.38
Liu et al. (2018c)	2018	Energy Convers. Manage.	405	57.86
Liu et al. (2018b)	2018	Energy Convers. Manage.	405	57.86
Hu et al. (2016)	2016	Renewable Energy	377	41.89
Chen et al. (2018)	2018	Energy Convers. Manage.	325	46.43
Qureshi et al. (2017)	2017	Applied Soft Computing	318	39.75
Liu et al. (2018a)	2018	Energy Convers. Manage.	273	39.00
Hong and Rioflorido (2019)	2019	Applied Energy	246	41.00
Jiang et al. (2021)	2021	Energy	161	40.25
Neshat et al. (2021)	2021	Energy Convers. Manage.	151	37.75
Wang et al. (2020b)	2020	IEEE Trans. Neural Netw. Learn. Syst.	134	26.80
Li et al. (2023)	2023	Applied Energy	126	63.00

4.6 Keyword Co-Occurrence and Thematic Evolution

The author-keyword co-occurrence network in Figure 4 identifies three central terms: wind power forecasting, deep learning and wind speed prediction. Their strongest linkages involve LSTM, Bi-LSTM, attention mechanisms, probabilistic forecasting, wavelet and mode decomposition, and hybrid learning. The network supports three broad thematic groupings: probabilistic and ensemble forecasting; decomposition-assisted hybrid DL; and attention-based or Transformer-oriented modeling. These cluster labels are interpretive summaries of network structure rather than mutually exclusive categories.

**Figure 4. Network visualization of author-defined keywords (at least three occurrences).**

The temporal overlay in Figure 5 indicates a progression from an earlier emphasis on numerical weather prediction, probabilistic methods and conventional time-series modeling toward recurrent and hybrid DL approaches, and more recently toward Bi-LSTM, attention mechanisms, transfer learning, VMD, ultra-short-term forecasting and federated learning. Because overlay color encodes average publication year rather than growth rate, the figure establishes relative recency but does not by itself identify the fastest-

growing cluster. The density map in Figure 6 confirms that deep learning, wind power forecasting and wind speed prediction form the field's thematic core, with renewable integration, wind farm optimization and real-time forecasting appearing as applied peripheral themes.

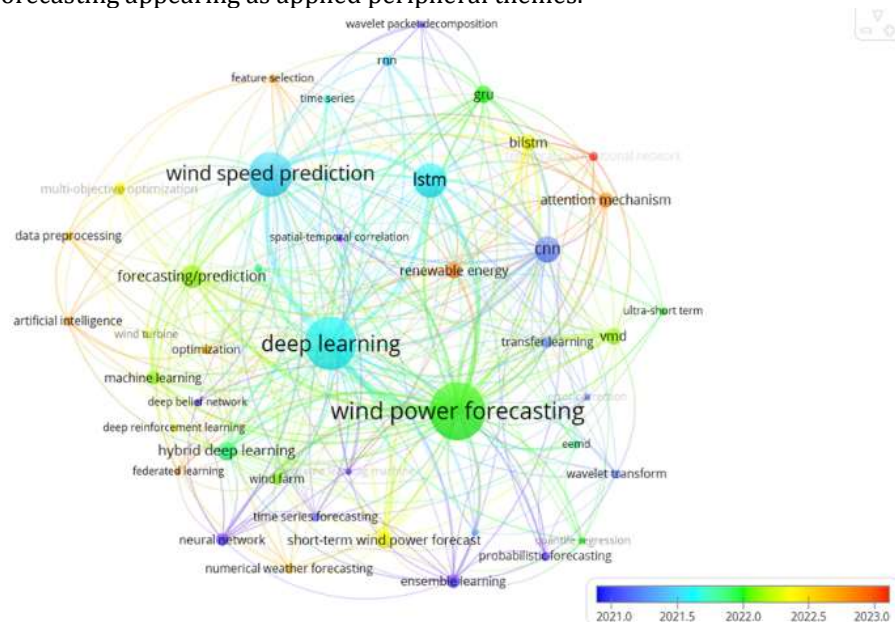


Figure 5. Temporal overlay visualization of author-defined keywords.

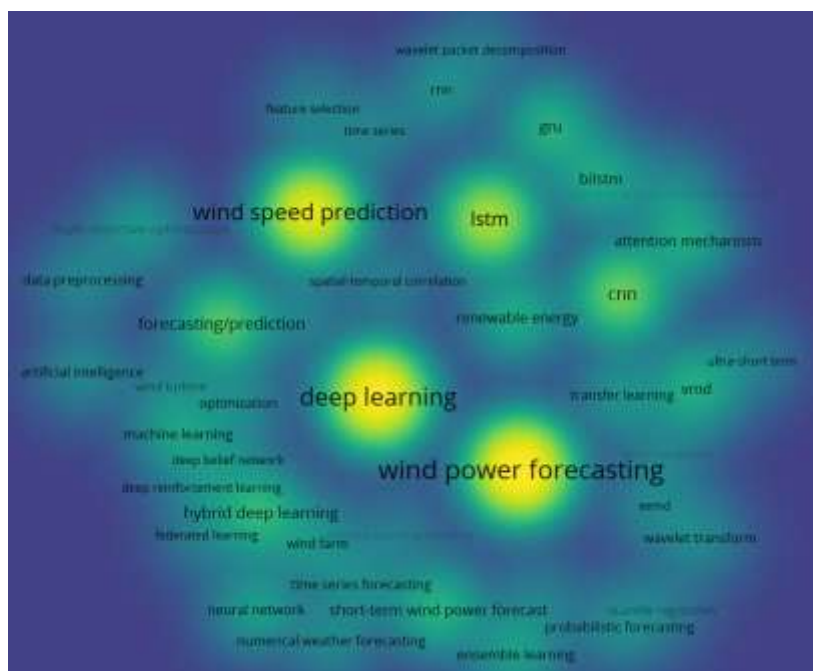


Figure 6. Density visualization of author-keyword co-occurrence intensity.

4.7 Synthesis and Research Priorities

The combined technical and bibliometric evidence shows rapid expansion alongside methodological convergence. Recurrent architectures, particularly LSTM and Bi-LSTM, remain central because they match the sequential structure of the problem and integrate readily with decomposition, attention and optimization modules. Hybridization is attractive because its components address different sources of difficulty: decomposition reduces non-stationarity, recurrent units' model temporal dependence, convolution captures local and spatial structure, attention reallocates model focus, and metaheuristics automate hyperparameter search. Each added component, however, increases computational cost, tuning burden, and the risk of data leakage or overfitting. Bibliometric prominence should not be equated with superior predictive performance, since reported accuracy also depends on dataset quality, horizon, train-

test design, baseline selection and metric choice. The rising visibility of attention and Transformer models therefore marks an emerging direction rather than a definitive displacement of recurrent architectures. Seven priorities follow from the identified gaps. First, physics-informed and constraint-aware models should improve physical consistency and extrapolation. Second, federated and privacy-preserving learning should support multi-owner wind-farm data. Third, transfer and domain-adaptation methods are needed for data-scarce and newly commissioned sites. Fourth, calibrated probabilistic forecasting should support operational risk assessment. Fifth, standardized open benchmark datasets require transparent chronological splits. Sixth, reproducible studies should report model size, inference latency, memory use, energy consumption, and statistical uncertainty. Seventh, benchmark datasets and validation studies should include underrepresented climatic regions where wind regimes, terrain, grid structure, and data availability differ from dominant research settings.

For practice, hybrid DL models are promising for short- and medium-term operation, but deployment readiness should be demonstrated rather than assumed. Operators require models that hold calibration through seasonal change, ramp events, curtailment, sensor faults and train-to-operation distribution shift, so operational studies should report rolling retraining, uncertainty intervals, fallback behavior, latency and robustness alongside average point error. Model selection should follow the decision context: reserve scheduling and market bidding benefit from probabilistic output, whereas turbine-level or edge deployment prioritizes compact models and low inference latency.

4.8 Limitations

Six limitations qualify these findings. The corpus is restricted to Scopus and predominantly English-language records, so regional or non-indexed work may be underrepresented. The title-only query improves precision at the expense of recall. Citation indicators favor older publications and are sensitive to database updates, self-citation, document type, and full-counting attribution. Author and institutional disambiguation reduces but does not eliminate metadata ambiguity. VOSviewer clusters depend on keyword normalization, occurrence thresholds, and algorithmic settings. Finally, the analysis maps metadata and keywords rather than the full text of all 324 publications, so thematic frequency is not a direct measure of model performance. Future work should conduct sensitivity analyses using Web of Science and IEEE Xplore, broader TITLE-ABS-KEY queries, fractional counting, citation normalization, and systematic full-text coding.

5. Conclusion

This review integrated a technical synthesis with bibliometric knowledge mapping to clarify how deep learning for wind energy forecasting has developed. Analysis of 324 Scopus-indexed publications from 2014 to 2024 shows rapid growth and a methodological core built on LSTM, Bi-LSTM, decomposition-assisted hybrids, probabilistic forecasting and attention mechanisms, with recent expansion toward Transformers, transfer and federated learning, explainable AI and physics-informed modeling. China leads in attributed output, while several lower-volume contributors achieve higher citation rates per paper, and international collaboration remains geographically uneven.

The principal contribution is not a ranking of individual algorithms but an integrated account of how methodological choices, publication structures and emerging themes interact. Progress now depends less on adding components to hybrid pipelines and more on evaluation discipline: standardized chronological benchmarks, calibrated uncertainty, multi-site and cross-climate validation, transparent ablation and statistical testing, and reporting of computational and deployment cost. These practices are necessary to move from promising benchmark results toward trustworthy forecasting systems for renewable energy operation and planning.

References

1. Abou Houran, M., Salman Bukhari, S. M., Zafar, M. H., Mansoor, M., & Chen, W. (2023). COA-CNN-LSTM: Coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications. *Appl. Energy*, 349, 121638. <https://doi.org/10.1016/j.apenergy.2023.121638>
2. Afrasiabi, M., Mohammadi, M., Rastegar, M., & Afrasiabi, S. (2021). Advanced Deep Learning Approach for Probabilistic Wind Speed Forecasting. *IEEE Trans. Ind. Informatics*, 17(1), 720–727. <https://doi.org/10.1109/TII.2020.3004436>
3. Alkhayat, G., & Mehmood, R. (2021). A review and taxonomy of wind and solar energy forecasting methods based on deep learning. *Energy AI*, 4, 100060. <https://doi.org/10.1016/j.egyai.2021.100060>
4. Aly, H. H. H. (2020). A novel deep learning intelligent clustered hybrid models for wind speed and

- power forecasting. *Energy*, 213, 118773. <https://doi.org/10.1016/j.energy.2020.118773>
5. Bandara, K., Hewamalage, H., Liu, Y. H., Kang, Y., & Bergmeir, C. (2021). Improving the accuracy of global forecasting models using time series data augmentation. *Pattern Recognit.*, 120, 108148. <https://doi.org/10.1016/j.patcog.2021.108148>
6. Bommidi, B. S., Teeparthi, K., & Kosana, V. (2023). Hybrid wind speed forecasting using ICEEMDAN and transformer model with novel loss function. *Energy*, 265, 126383. <https://doi.org/10.1016/j.energy.2022.126383>
7. Chen, J., Zeng, G., Zhou, W., Du, W., & Lu, K. (2018). Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. *Energy Convers. Manag.*, 165, 681–695. <https://doi.org/10.1016/j.enconman.2018.03.098>
8. Ding, Y., Chen, Z., Zhang, H., Wang, X., & Guo, Y. (2022). A short-term wind power prediction model based on CEEMD and WOA-KELM. *Renew. Energy*, 189, 188–198. <https://doi.org/10.1016/j.renene.2022.02.108>
9. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
10. Faruque, M. O., Hossain, M. A., Islam, M. R., Alam, S. M. M., & Karmaker, A. K. (2024). Very short-term wind power forecasting for real-time operation using hybrid deep learning model with optimization algorithm. *Clean. Energy Syst.*, 9, 100129. <https://doi.org/10.1016/j.cles.2024.100129>
11. Gu, B., Hu, H., Zhao, J., Zhang, H., & Liu, X. (2023). Short-term wind power forecasting and uncertainty analysis based on FCM-WOA-ELM-GMM. *Energy Reports*, 9, 807–819. <https://doi.org/10.1016/j.egyr.2022.11.202>
12. Harbola, S., & Coors, V. (2019). One dimensional convolutional neural network architectures for wind prediction. *Energy Conversion and Management*. <https://www.sciencedirect.com/science/article/pii/S0196890419305527>
13. Hong, Y. Y., & Rioflorido, C. L. P. P. (2019). A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. *Appl. Energy*, 250, 530–539. <https://doi.org/10.1016/j.apenergy.2019.05.044>
14. Hong, Y. Y., & Satriani, T. R. A. (2020). Day-ahead spatiotemporal wind speed forecasting using robust design-based deep learning neural network. *Energy*, 209, 118441. <https://doi.org/10.1016/j.energy.2020.118441>
15. Hossain, M. A., Chakraborty, R. K., Elsayah, S., & Ryan, M. J. (2021). Very short-term forecasting of wind power generation using hybrid deep learning model. *J. Clean. Prod.*, 296, 126564. <https://doi.org/10.1016/j.jclepro.2021.126564>
16. Hu, Q., Zhang, R., & Zhou, Y. (2016). Transfer learning for short-term wind speed prediction with deep neural networks. *Renew. Energy*, 85, 83–95. <https://doi.org/10.1016/j.renene.2015.06.034>
17. Huangfu, J., Shi, Z., & Liu, Y. (2022). Short-term Forecast of Wind Speed Based on Hybrid Multi-Stage Decomposition and LSTM Optimized by an Improved WOA. *Proc. - 2022 8th Int. Conf. Big Data Inf. Anal. BigDIA 2022*, 193–198. <https://doi.org/10.1109/BigDIA56350.2022.9874126>
18. Ibrahim, A., et al. (2021). Wind Speed Ensemble Forecasting Based on Deep Learning Using Adaptive Dynamic Optimization Algorithm. *IEEE Access*, 9, 125787–125804. <https://doi.org/10.1109/ACCESS.2021.3111408>
19. Jiang, P., Liu, Z., Niu, X., & Zhang, L. (2021). A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting. *Energy*, 217, 119361. <https://doi.org/10.1016/j.energy.2020.119361>
20. Kosana, V., Teeparthi, K., & Madasthu, S. (2022). Hybrid wind speed prediction framework using data pre-processing strategy based autoencoder network. *Electr. Power Syst. Res.*, 206, 107821. <https://doi.org/10.1016/j.epsr.2022.107821>
21. Krishna Rayi, V., Mishra, S. P., Naik, J., & Dash, P. K. (2022). Adaptive VMD based optimized deep learning mixed kernel ELM autoencoder for single and multistep wind power forecasting. *Energy*, 244, 122585. <https://doi.org/10.1016/j.energy.2021.122585>
22. Lagos, A., et al. (2022). State-of-the-Art Using Bibliometric Analysis of Wind-Speed and -Power Forecasting Methods Applied in Power Systems. *Energies*, 15(18). <https://doi.org/10.3390/en15186545>
23. Li, C., Tang, G., Xue, X., Saeed, A., & Hu, X. (2020). Short-Term Wind Speed Interval Prediction Based on Ensemble GRU Model. *IEEE Trans. Sustain. Energy*, 11(3), 1370–1380. <https://doi.org/10.1109/TSTE.2019.2926147>
24. Li, Y., Wang, R., Li, Y., Zhang, M., & Long, C. (2023). Wind power forecasting considering data privacy

- protection: A federated deep reinforcement learning approach. *Appl. Energy*, 329, 120291. <https://doi.org/10.1016/j.apenergy.2022.120291>
25. Lim, W. M., Kumar, S., & Donthu, N. (2024). How to combine and clean bibliometric data and use bibliometric tools synergistically: Guidelines using metaverse research. *J. Bus. Res.*, 182, 114760. <https://doi.org/10.1016/j.jbusres.2024.114760>
26. Lin, Z., Liu, X., & Collu, M. (2020). Wind power prediction based on high-frequency SCADA data along with isolation forest and deep learning neural networks. *Int. J. Electr. Power Energy Syst.*, 118, 105835. <https://doi.org/10.1016/j.ijepes.2020.105835>
27. Liu, H., Mi, X., & Li, Y. (2018a). Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. *Energy Convers. Manag.*, 166, 120–131. <https://doi.org/10.1016/j.enconman.2018.04.021>
28. Liu, H., Mi, X., & Li, Y. (2018b). Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy Convers. Manag.*, 159, 54–64. <https://doi.org/10.1016/j.enconman.2018.01.010>
29. Liu, H., Mi, X. W., & Li, Y. F. (2018c). Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. *Energy Convers. Manag.*, 156, 498–514. <https://doi.org/10.1016/j.enconman.2017.11.053>
30. Liu, H., Mi, X., Li, Y., Duan, Z., & Xu, Y. (2019). Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional Gated Recurrent Unit network and Support Vector Regression. *Renew. Energy*, 143, 842–854. <https://doi.org/10.1016/j.renene.2019.05.039>
31. Liu, H., Yu, C., Wu, H., Duan, Z., & Yan, G. (2020a). A new hybrid ensemble deep reinforcement learning model for wind speed short term forecasting. *Energy*, 202, 117794. <https://doi.org/10.1016/j.energy.2020.117794>
32. Liu, B., Zhao, S., Yu, X., Zhang, L., & Wang, Q. (2020b). A novel deep learning approach for wind power forecasting based on WD-LSTM model. *Energies*, 13(18), 1–17. <https://doi.org/10.3390/en13184964>
33. Liu, Y., et al. (2020c). Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model. *Appl. Energy*, 260, 114259. <https://doi.org/10.1016/j.apenergy.2019.114259>
34. Liu, Z., Hara, R., & Kita, H. (2021). Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting. *Energy Convers. Manag.*, 238, 114136. <https://doi.org/10.1016/j.enconman.2021.114136>
35. Lv, S. X., & Wang, L. (2023). Multivariate wind speed forecasting based on multi-objective feature selection approach and hybrid deep learning model. *Energy*, 263, 126100. <https://doi.org/10.1016/j.energy.2022.126100>
36. Ma, Z., et al. (2020). Application of hybrid model based on double decomposition, error correction and deep learning in short-term wind speed prediction. *Energy Convers. Manag.*, 205, 112345. <https://doi.org/10.1016/j.enconman.2019.112345>
37. Ma, Z., & Mei, G. (2022). A hybrid attention-based deep learning approach for wind power prediction. *Applied Energy*. <https://www.sciencedirect.com/science/article/pii/S0306261922009138>
38. Mbugua, J. M., & Hiraga, Y. (2025). Recent Advances in Long-Term Wind-Speed and -Power Forecasting: A Review. *Climate*, 13(8), 1–30. <https://doi.org/10.3390/cli13080155>
39. Meng, A., et al. (2022). A hybrid deep learning architecture for wind power prediction based on bi-attention mechanism and crisscross optimization. *Energy*. <https://www.sciencedirect.com/science/article/pii/S0360544221020430>
40. Mujeeb, S., Alghamdi, T. A., Ullah, S., Fatima, A., Javaid, N., & Saba, T. (2019). Exploiting deep learning for wind power forecasting based on big data analytics. *Appl. Sci.*, 9(20). <https://doi.org/10.3390/app9204417>
41. Neshat, M., et al. (2021). A deep learning-based evolutionary model for short-term wind speed forecasting : A case study of the Lillgrund offshore wind farm. *Energy Convers. Manag.*, 236, 114002. <https://doi.org/10.1016/j.enconman.2021.114002>
42. Niu, D., Sun, L., Yu, M., & Wang, K. (2022). Point and interval forecasting of ultra-short-term wind power based on a data-driven method and hybrid deep learning model. *Energy*, 254, 124384. <https://doi.org/10.1016/j.energy.2022.124384>
43. Peng, Z., et al. (2020). A novel deep learning ensemble model with data denoising for short-term

- wind speed forecasting. *Energy Convers. Manag.*, 207, 112524.
<https://doi.org/10.1016/j.enconman.2020.112524>
44. Peng, X., et al. (2021). EALSTM-QR: Interval wind-power prediction model based on numerical weather prediction and deep learning. *Energy*.
<https://www.sciencedirect.com/science/article/pii/S0360544220327997>
45. Qureshi, A. S., Khan, A., Zameer, A., & Usman, A. (2017). Wind power prediction using deep neural network based meta regression and transfer learning. *Appl. Soft Comput. J.*, 58, 742–755.
<https://doi.org/10.1016/j.asoc.2017.05.031>
46. Rajaperumal, T. A., & Christopher Columbus, C. (2025). Enhanced wind power forecasting using machine learning, deep learning models and ensemble integration. *Sci. Rep.*, 15(1), 1–22.
<https://doi.org/10.1038/s41598-025-05250-3>
47. Samadianfard, S., et al. (2020). Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm. *Energy Reports*, 6, 1147–1159.
<https://doi.org/10.1016/j.egyr.2020.05.001>
48. Subih, A. A., Pasupuleti, J., & Paw, J. K. S. (2026). Enhanced wind speed prediction with deep learning models optimized by metaheuristic algorithms. *AIP Conf. Proc.*, 3407(1).
<https://doi.org/10.1063/5.0318629>
49. Tian, Z. (2021). Modes decomposition forecasting approach for ultra-short-term wind speed. *Appl. Soft Comput.*, 105, 107303. <https://doi.org/10.1016/j.asoc.2021.107303>
50. Verdone, A., Panella, M., De Santis, E., & Rizzi, A. (2025). A review of solar and wind energy forecasting: From single-site to multi-site paradigm. *Appl. Energy*, 392, 126016.
<https://doi.org/10.1016/j.apenergy.2025.126016>
51. Wang, H. Z., Li, G. Q., Wang, G. B., Peng, J. C., Jiang, H., & Liu, Y. T. (2017). Deep learning based ensemble approach for probabilistic wind power forecasting. *Appl. Energy*, 188, 56–70.
<https://doi.org/10.1016/j.apenergy.2016.11.111>
52. Wang, J., & Li, Y. (2018). Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy. *Appl. Energy*, 230, 429–443. <https://doi.org/10.1016/j.apenergy.2018.08.114>
53. Wang, X., Yu, Q., & Yang, Y. (2018). Short-term wind speed forecasting using variational mode decomposition and support vector regression. *J. Intell. Fuzzy Syst.*, 34(6), 3811–3820.
<https://doi.org/10.3233/JIFS-169553>
54. Wang, R., Chen, B., Liu, B., Lin, H., & Chen, W. (2019). The Application Research of Deep Learning Based Method for Short-term Wind Speed Forecasting. *Proc. 2019 IEEE 4th Adv. Inf. Technol. Electron. Autom. Control Conf. IAEAC 2019*, 2378–2381.
<https://doi.org/10.1109/IAEAC47372.2019.8997793>
55. Wang, J., Niu, T., Lu, H., Yang, W., & Du, P. (2020a). A Novel Framework of Reservoir Computing for Deterministic and Probabilistic Wind Power Forecasting. *IEEE Trans. Sustain. Energy*, 11(1), 337–349. <https://doi.org/10.1109/TSTE.2019.2890875>
56. Wang, R., Li, C., Fu, W., & Tang, G. (2020b). Deep Learning Method Based on Gated Recurrent Unit and Variational Mode Decomposition for Short-Term Wind Power Interval Prediction. *IEEE Trans. Neural Networks Learn. Syst.*, 31(10), 3814–3827.
<https://doi.org/10.1109/TNNLS.2019.2946414>
57. Wang, S., Wang, J., Lu, H., & Zhao, W. (2021a). A novel combined model for wind speed prediction – Combination of linear model, shallow neural networks, and deep learning approaches. *Energy*, 234, 121275. <https://doi.org/10.1016/j.energy.2021.121275>
58. Wang, L., Tao, R., Hu, H., & Zeng, Y. R. (2021b). Effective wind power prediction using novel deep learning network: Stacked independently recurrent autoencoder. *Renew. Energy*, 164, 642–655.
<https://doi.org/10.1016/j.renene.2020.09.108>
59. Wu, W., Chen, K., Qiao, Y., & Lu, Z. (2016). Probabilistic Short-term Wind Power Forecasting Based on Deep Neural Networks. <https://doi.org/10.1109/PMAPS.2016.7764155>
60. Wu, Z., Luo, G., Yang, Z., Guo, Y., Li, K., & Xue, Y. (2022). A comprehensive review on deep learning approaches in wind forecasting applications. *CAAI Trans. Intell. Technol.*, 7(2), 129–143.
<https://doi.org/10.1049/cit2.12076>
61. Xiang, L., Li, J., Hu, A., & Zhang, Y. (2020). Deterministic and probabilistic multi-step forecasting for short-term wind speed based on secondary decomposition and a deep learning method. *Energy Convers. Manag.*, 220, 113098. <https://doi.org/10.1016/j.enconman.2020.113098>
62. Xiong, B., Lou, L., Meng, X., Wang, X., Ma, H., & Wang, Z. (2022). Short-term wind power forecasting based on Attention Mechanism and Deep Learning. *Electr. Power Syst. Res.*, 206, 107776.
<https://doi.org/10.1016/j.epsr.2022.107776>

63. Xiong, J., Peng, T., Tao, Z., Zhang, C., Song, S., & Nazir, M. S. (2023). A dual-scale deep learning model based on ELM-BiLSTM and improved reptile search algorithm for wind power prediction. *Energy*, 266, 126419. <https://doi.org/10.1016/j.energy.2022.126419>
64. Yan, X., Liu, Y., Xu, Y., & Jia, M. (2020). Multistep forecasting for diurnal wind speed based on hybrid deep learning model with improved singular spectrum decomposition. *Energy Convers. Manag.*, 225, 113456. <https://doi.org/10.1016/j.enconman.2020.113456>
65. Yang, W., Wang, J., Lu, H., Niu, T., & Du, P. (2019). Hybrid wind energy forecasting and analysis system based on divide and conquer scheme: A case study in China. *J. Clean. Prod.*, 222, 942–959. <https://doi.org/10.1016/j.jclepro.2019.03.036>
66. Ye, P., Li, Y., Zhang, H., & Shen, H. (2020). Bibliometric analysis on the research of offshore wind power based on web of science. *Econ. Res. Istraz.*, 33(1), 887–903. <https://doi.org/10.1080/1331677X.2020.1734853>
67. Yin, H., Ou, Z., Huang, S., & Meng, A. (2019). A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition. *Energy*, 189, 116316. <https://doi.org/10.1016/j.energy.2019.116316>
68. Yu, C., Li, Y., Xiang, H., & Zhang, M. (2018). Data mining-assisted short-term wind speed forecasting by wavelet packet decomposition and Elman neural network. *J. Wind Eng. Ind. Aerodyn.*, 175, 136–143. <https://doi.org/10.1016/j.jweia.2018.01.020>
69. Yu, R., et al. (2019). Scene learning: Deep convolutional networks for wind power prediction by embedding turbines into grid space. *Appl. Energy*, 238, 249–257. <https://doi.org/10.1016/j.apenergy.2019.01.010>
70. Zaidi, A. (2024). Utilisation of Deep Learning (DL) and Neural Networks (NN) Algorithms for Energy Power Generation: A Social Network and Bibliometric Analysis (2004-2022). *Int. J. Energy Econ. Policy*, 14(1), 172–183. <https://doi.org/10.32479/ijee.15161>
71. Zhang, S., Chen, Y., Xiao, J., Zhang, W., & Feng, R. (2021). Hybrid wind speed forecasting model based on multivariate data secondary decomposition approach and deep learning algorithm with attention mechanism. *Renewable Energy*, 174, 688–704. <https://doi.org/10.1016/j.renene.2021.04.091>
72. Zhang, C., Ma, H., Hua, L., Sun, W., Nazir, M. S., & Peng, T. (2022a). An evolutionary deep learning model based on TVFEMD, improved sine cosine algorithm, CNN and BiLSTM for wind speed prediction. *Energy*, 254, 124250. <https://doi.org/10.1016/j.energy.2022.124250>
73. Zhang, L., Ling, J., & Lin, M. (2022b). Artificial intelligence in renewable energy: A comprehensive bibliometric analysis. *Energy Reports*, 8, 14072–14088. <https://doi.org/10.1016/j.egyr.2022.10.347>
74. Zhang, W., Lin, Z., & Liu, X. (2022c). Short-term offshore wind power forecasting - A hybrid model based on Discrete Wavelet Transform (DWT), Seasonal Autoregressive Integrated Moving Average (SARIMA), and deep-learning-based Long Short-Term Memory (LSTM). *Renew. Energy*, 185, 611–628. <https://doi.org/10.1016/j.renene.2021.12.100>
75. Zhu, Q., Chen, J., Zhu, L., Duan, X., & Liu, Y. (2018). Wind speed prediction with spatio-temporal correlation: A deep learning approach. *Energies*. <https://www.mdpi.com/274622>