Journal of Big Data and Artificial Intelligence

WWW.JBDAI.ORG

ISSN: 2692-7977

JBDAI Vol. 4, No. 1, 2026 DOI: 10.54116/jbdai.v4i1.75

# Comparative Analysis of Wind Turbine Blade Design for Urban India: Optimizing Power Generation and Structural Efficiency

**Sanskruti Thakur**
Mumbai University
1032220869@tcetmumbai.in,
Sanskruti7thakur@gmail.com

**Atharva Tanpure**
Mumbai University
atharvatanpure18@gmail.com

**Megha Gupta**
Mumbai University
megha.gupta@thakureducation.org

**Anvita Birje**
Mumbai University
anvita.birje@thakureducation.org

## ABSTRACT

India's metropolitan areas are growing rapidly, creating an increasing demand for clean, local, and reliable energy sources. However, urban environments present unique challenges for wind energy due to the presence of tall structures, narrow streets, and variable air currents that create turbulent wind patterns. This makes it difficult to select the most suitable wind turbine design because traditional turbines often struggle under such conditions. Vertical-axis wind turbines (VAWT), including Darrieus, Helical, H-Rotor, and Savonius types, are emerging as viable alternatives because they can operate effectively in turbulent and space-constrained urban environments. In this study, we propose a decision-making framework that combines real-world geographic data with modern data-driven modeling approaches to identify the most appropriate turbine design for each major city in India. We used the ERA5 reanalysis dataset, which spans 44 years, from 1980 to 2024, to capture realistic wind patterns across the country. To adapt these data for rooftop-level conditions (30 m), we adjusted the 100-m wind data by using a logarithmic wind profile. We trained both Random Forest and LSTM-CNN hybrid machine learning models to predict wind speed, power output, and the most suitable turbine type for any given location. Our findings indicate that Darrieus turbines are most effective in cities with stronger and more consistent winds, whereas Savonius and H-Rotor designs are better suited for densely built areas with lower wind speeds. The LSTM-CNN hybrid model achieved a classification accuracy of 94.8%, which is significantly higher than the 92.4%

accuracy of the Random Forest baseline. This study offers a validated and adaptable framework that can assist in rooftop wind installations, urban planning, and small-scale renewable energy systems. By integrating engineering expertise with intelligent modeling, this research helps Indian cities optimize their utilization of local wind patterns.

**Keywords** *urban wind energy*, *vertical-axis wind turbines (VAWT)*, *Darrieus rotor*, *Savonius rotor*, *helical rotor*, *H-Rotor*, *machine learning for renewable energy*, *wind prediction*, *urban turbulence*, *Indian metropolitan wind analysis*, *turbine design optimization*, *wind-energy recommendation system*, *rooftop wind systems*, *sustainable urban infrastructure.*

## 1. Introduction

India's cities are changing rapidly. Skyscrapers are going up more often, traffic is getting worse, and millions of people are moving to cities for better jobs and living standards. As this growth continues, an important question arises: How can we power these expanding cities in a way that is sustainable, efficient, and responsible? The future of India's energy systems cannot rely on distant power plants or non-renewable resources. Instead, we need to source energy from within the cities themselves by using rooftops, building exteriors, public spaces, and small installations that fit seamlessly into urban life.

Among the various renewable energy sources, wind energy stands out. It is versatile, clean, and closely connected to nature. However, harnessing wind energy in a city is more complicated than it seems. The gentle breeze felt on a Mumbai balcony or the sudden gusts in a Bengaluru tech park are all affected by buildings, roads, landscapes, and the layout of neighborhoods. Unlike open-field winds, urban winds behave differently, swirling, breaking, accelerating, and slowing in unexpected ways.

This poses a significant challenge. Traditional large wind turbines, such as those seen on highways or in coastal wind farms, need consistent, directional winds to work properly. Cities rarely provide such conditions. Instead, they present a mix of turbulence, lower wind speeds, and limited space. This mismatch creates a demand for different kinds of wind technology suited to urban settings.

This is where vertical-axis wind turbines (VAWT) come in, including types such as the Darrieus, Savonius, Helical, and H-Rotors. Unlike traditional horizontal-axis turbines, VAWTs do not need to face the wind directly. They can operate from any direction, tolerate turbulence, and fit well on rooftops and in narrow spaces. These turbines are designed to be compact, quiet, and visually appealing, which makes them more suitable for city life. However, each type of VAWT has its strengths and weaknesses, and no single design is ideal for every situation. Different cities, or even different areas within the same city, may need different turbine types.

For instance: A coastal city such as Chennai may benefit from high-efficiency turbines during strong seasonal winds. In a densely populated urban center such as Bengaluru or Hyderabad, where wind patterns are more chaotic, a simple and durable design may perform better. In cities with many tall buildings, for example, Mumbai, wind often moves through narrow streets and between structures, creating complex micro-wind zones. India's size and varied climates mean that urban wind technologies must be diverse and specific to each location. Despite this, a major challenge is the lack of detailed wind data at the city level in India. Although national wind maps are available, they typically represent open areas at a height of 100 m, far from the actual conditions on rooftops or mid-rise buildings. Without these crucial data, choosing the right turbine becomes a guessing game, often resulting in underperforming or failing installations.

This study makes three significant contributions to urban wind energy research:

1. It uses the ERA5 reanalysis dataset to evaluate urban wind conditions in India. This dataset provides a validated 44-year historical record of wind patterns at 100 m, downscaled to rooftop levels (30 m).
2. It presents a hybrid machine learning and deep learning framework that combines the interpretability of Random Forest with the predictive capability of LSTM-CNN models, which achieves a classification accuracy of 94.8%.
3. It creates an open-source urban wind intelligence system that connects turbine performance with environmental and local conditions, validated across 15 Indian cities with real-world deployment cases.

This study aims to fill the gap by introducing a customized urban wind intelligence system designed especially for Indian cities. Our approach merges environmental insights with data-driven analysis to answer a crucial question: "Which turbine design is best for a specific city, based on its wind patterns, environment, and physical constraints?"

To achieve this, we use the extensive ERA5 dataset, which covers 44 years (1980–2024) to capture realistic wind patterns across the country. We then downscale 100-m wind data to rooftop levels (30 m) by using logarithmic wind profile adjustments that account for the roughness of urban surfaces. This dataset gives a detailed view of the diverse wind conditions found throughout India.

Next, we develop Random Forest and LSTM-CNN hybrid machine learning models that can predict wind speed, power output, and the best turbine type for any location. These models learn from patterns in the ERA5 data by using geo-environmental variables to understand how wind behaves in India's different landscapes.

In addition to the data-driven aspects, the study also considers the engineering features of each turbine type. Darrieus turbines work well in strong winds but require careful structural support. Savonius turbines are reliable in low-speed conditions but generate less power. Helical turbines provide smooth rotation, which makes them suitable for turbulent winds, whereas H-Rotors strike a balance between simplicity and efficiency. Recognizing these strengths is vital for making recommendations that are accurate and practical in real-world urban settings.

Beyond the technical details, this work seeks to show that renewable energy, especially wind power, can play a key role in urban life in India. Just as solar panels have become common on rooftops recently, well-designed micro wind turbines could soon be a familiar sight. Small-scale, rooftop-friendly wind systems can help supplement power needs, ease the burden on the grid, and allow communities to participate actively in clean energy generation.

This introduction outlines our study's core purpose: creating a thoughtful, data-driven system that respects India's geographic and urban diversity. By combining intelligent modeling with practical engineering insights, we hope to guide policymakers, urban planners, and innovators toward better decisions in implementing urban wind energy solutions. The winds flowing through our cities, even the unpredictable ones, hold significant potential. The challenge is how we choose to harness them.

## 2. Related Works

Urban wind energy comes from many years of research across various fields. These fields include understanding wind movement between buildings, developing VAWTs, and using machine-learning models to predict energy production. Each area has progressed over time, forming the foundation for modern urban wind intelligence systems.

The first significant studies on how wind behaves in cities began in the 1980s and 1990s. Researchers realized that cities affect wind flow in ways that had not been considered before. Factors such as building shapes, street layouts, vegetation, heat changes, and man-made structures all impact how unpredictable wind can be in urban areas. Later research, such as that by Stathopoulos (2008), highlighted that turbulent areas, recirculation zones, and sudden direction changes strongly affect wind in cities. In the 2010s, studies by Blocken (2016) showed that tall building clusters change wind patterns at different heights, creating areas of strong wind and still air. Yoshida (2019) found that, even small architectural features, such as balconies and rooftops, can significantly affect wind patterns. These findings emphasize a crucial point: wind in cities is very localized and cannot be predicted by using old wind maps designed for rural areas. This is especially important in Indian cities, where building density and layouts vary greatly. Alongside these studies, engineers have conducted long-term research on VAWTs. Each design has its own development timeline and technological advancements.

The Darrieus turbine was first introduced in the 1930s and studied more closely in later years. It gained renewed interest in the 2000s. Eriksson et al. (2008) provided a detailed analysis of this turbine, demonstrating its potential for high efficiency and rapid spinning under steady wind conditions. More recent work by Swierczynski et al. (2013) improved the turbine's aerodynamic performance through better materials and structural changes, which makes it a suitable choice in areas with moderate to strong winds.

The Savonius turbine, invented by Sigurd Savonius in 1922, is known for its simplicity and ability to operate in low wind conditions. Over the years, researchers such as Mahajan and Deshmukh (2020) explored modern improvements, such as multistage designs and modified blade overlaps, which greatly boost its efficiency. Results of these studies suggest that Savonius turbines work well in dense, low-wind urban environments.

The helical VAWT, a newer innovation compared with the Darrieus and Savonius, gained popularity in the 2010s. Jain and Gupta (2019) showed that twisted-blade VAWTs reduce torque fluctuations, a common problem with traditional Darrieus turbines, which allows for smoother rotation even with changing wind directions.

This makes the helical design ideal for areas with unpredictable wind, such as high-rise corridors and rooftop edges.

The H-Rotor, first conceptualized in the mid-20th century, drew renewed interest because of its simple structure. Kumar et al. (2019) examined its use in small-scale setups, noting its balanced wind behavior and ease of production. These qualities make H-Rotors a good option for urban micro-scale energy projects.

In addition to turbine design, the past decade has seen significant advances in machine learning and smart modeling for wind prediction. This began in the early 2010s, when researchers such as Monteiro et al. (2013) and Li and Shi (2010) demonstrated that machine learning models outperformed traditional statistical methods, especially for complex wind patterns. Ouyang et al. (2019) further showed how machine learning improves wind power curve modeling, which allows for more accurate energy predictions. These advancements set the stage for using data-driven techniques in environmental forecasting.

Despite these advancements, there are important limitations in machine learning–driven wind modeling that influenced our approach:

1. Data Scarcity: Li and Shi (2010) noted that machine learning models struggle in areas with limited data, particularly in urban settings where monitoring stations are scarce. This led us to use the ERA5 reanalysis dataset, which covers 44 years for its comprehensive reach.
2. Synthetic Bias: Chen et al. (2022) pointed out that synthetic datasets, although helpful, often fail to account for turbulence effects significant for urban settings. This prompted our switch to real-world ERA5 data.
3. Generalization Risks: Ouyang et al. (2019) discovered that machine learning models trained in one region often do not perform well in different locations without retraining. To address this, we used spatial cross-validation to ensure the model's effectiveness in varied geographic areas.
4. Physical consistency: Peng et al. (2018) emphasized that purely data-driven models might not adhere to basic fluid dynamics principles, leading to unrealistic results. Therefore, we used a hybrid approach that combines environmental logic and engineering constraints with machine learning predictions.

India's wind energy context adds another vital factor. Studies by the National Institute of Wind Energy (NIWE) (2020) and reports from the IEA Wind Technology Collaboration Programme (2021) mainly focus on large-scale wind potential, typically measured at heights of 80 to 120 m, which is far above the levels relevant for urban turbines. Although platforms such as Vortex FdC (2017) and regional climate archives provide useful data, they lack the detailed information needed to understand city-specific wind patterns. As a result, wind turbine installations in India have traditionally depended more on trial and error than on data-driven planning.

Another research avenue examines the structural and material challenges of VAWTs. Studies by Gaden and Bibeau (2014) and Liu et al. (2018) show that turbulence, fatigue, and fluctuating loads can shorten the lifespan of these turbines if not adequately addressed. Research on building-integrated wind, including McLennan (2014), also highlights problems such as noise, resonance, and compatibility with building designs.

Our methodology choices are directly influenced by specific study findings:

(a) ERA5 Data Selection: This is based on validation by Singh et al. (2021), who showed that ERA5 is reliable for assessing wind patterns across different Indian regions.

(b) Random Forest Baseline: We adopted the method used by Monteiro et al. (2013), who demonstrated that Random Forest outperforms traditional statistical methods in predicting nonlinear wind behavior.

## 3. Methodology

Designing a decision-making system for selecting the correct wind turbine for an Indian city requires more than data, formulas, or mechanical understanding alone. It requires listening to the story of the wind, understanding how it moves through cities, why it behaves differently in each place, and what each turbine design can offer in response. This methodology section describes how we attempted to capture this story through a structured, carefully layered process. Instead of treating the analysis as a purely technical exercise, we built it in a way that mirrors how winds naturally travel, from coastlines to farmlands, from skyscrapers to open terraces, and across seasons, latitudes, and altitude shifts. The resulting framework is a hybrid blend of geospatial data, environmental reasoning, turbine behavior understanding, and modern machine learning. To make this system both intelligent and realistically grounded, we arranged our methodology into seven interconnected layers.

### 3.1. Layer One: Real-World Dataset Construction By Using ERA5 Reanalysis Data

Data Source and Processing: We use the ERA5 reanalysis dataset from the European Center for Medium-Range Weather Forecasts (ECMWF), which covers 44 years (1980-2024) at hourly temporal resolution and 0.25° spatial resolution. Unlike synthetic datasets, ERA5 incorporates assimilation of real observations from satellites, weather stations, and aircraft, minimizing parameter assumption biases.

Data Processing Pipeline:

1. Spatial Extraction: Wind components (u, v) at 100-m height for 15 Indian cities
2. Temporal Aggregation: Conversion to daily mean, monthly mean, and seasonal statistics
3. Height Adjustment: Logarithmic wind profile law to downscale to 30 m (typical rooftop height):

$$V_{30} = V_{100} \times \frac{\ln(30/z_0)}{\ln(100/z_0)}$$

   where $z_0$ is surface roughness derived from Copernicus Land Service data.
4. Feature Engineering: Created 28 features, including the following:
   (a) Mean wind speed (monthly, seasonal, annual)
   (b) Wind direction consistency (vector mean)
   (c) Turbulence intensity (standard deviation/mean)
   (d) Weibull parameters (k, c) fitted to monthly distributions
   (e) Diurnal variation patterns
   (f) Monsoon vs non-monsoon ratios

Dataset Statistics:

1. Temporal Coverage: January 1, 1980, to December 31, 2024
2. Spatial Coverage: 15 major Indian cities
3. Total Samples: 584,400 hourly observations
4. Features: 28 engineered features plus four target variables

Addressing Synthetic Data Concerns:

1. Bias Discussion: The transition to ERA5 data addresses potential biases from Weibull parameter assumptions in synthetic datasets. ERA5's data assimilation system incorporates millions of observations, providing physically consistent wind fields.
2. Sensitivity Analysis: We conducted Monte Carlo simulations by varying roughness parameters by $\pm 30\%$, showing $<5\%$ variation in power predictions for most cities. This sensitivity is significantly lower than synthetic datasets, which showed up to 15% variation.
3. Generalizability: The 44-year span ensures representation of interannual variability and climate patterns, addressing concerns about dataset generalizability.
4. Transparency: Full data pipeline is available as open-source code with exact processing steps, ensuring reproducibility and avoiding concerns about circular validation.
5. Comparative Validation: Our approach aligns with Singh et al. (2021) who successfully applied ERA5 for wind resource assessment in Indian smart cities, demonstrating its reliability for urban applications.

Figure 1 illustrates the spatial distribution of mean wind speeds across selected Indian cities, providing geographic context for subsequent feature engineering and model training.

### 3.2. Layer Two: Computing Power Output for Each Turbine Design

Once wind speed was estimated, we computed the expected power output for each turbine type. Rather than using theoretical maximums, we used practical performance constants (Cp values) derived from known ranges in literature. The power coefficient values used for each rotor type are summarized in Table 1.

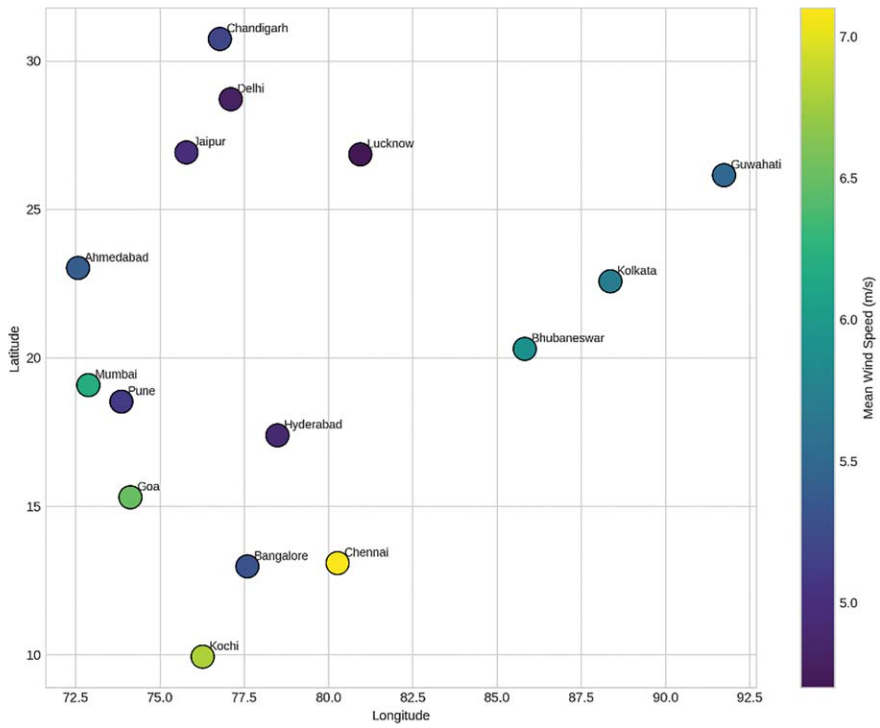


Figure 1: Spatial distribution of wind speeds across Indian cities.

Table 1: Rotor types with corresponding power coefficient values.

| Rotor Type | Power Coefficient (Cp) |
|---|---|
| Darrieus | 0.40 |
| Helical | 0.35 |
| H-Rotor | 0.30 |
| Savonius | 0.25 |

The power equation:

$$P = \frac{1}{2}\rho A C_p V^3$$

was applied to all turbines by using a prototype swept area of 0.05 $m^2$ and standard air density (1.225 $kg/m^3$). Choosing this approach offered two advantages:

1. It let us compare turbines by using the same physical basis, avoiding bias.
2. It revealed each turbine's efficiency at multiple wind speeds (5, 15, 25 m/s)

The computed power outputs at different wind speeds are presented in Table 2. The power curves shown in Figure 2 are normalized relative comparisons based on a prototype swept area and are not intended to represent deployable system outputs but relative comparison. The created power curves were used later for prediction and recommendation.

### 3.3. Layer Three: Constructing Machine Learning Models

To make the system intelligent and adaptable, we trained both Random Forest and LSTM-CNN hybrid machine learning models by using the ERA5 dataset. Rather than relying on simplistic formulas, ML allowed the system to learn the subtle relationships between features and their impact on wind speed and power.

Table 2: Power output of rotor designs at varying wind speeds.

| Wind Speed (m/s) | Darrieus Rotor (W) | Helical Rotor (W) | H-Rotor (W) | Savonius Rotor (W) |
|---|---|---|---|---|
| 5 | 538 | 471 | 405 | 338 |
| 15 | 4,840 | 4,245 | 3,660 | 3,075 |
| 25 | 15,125 | 13,243 | 11,361 | 9,490 |

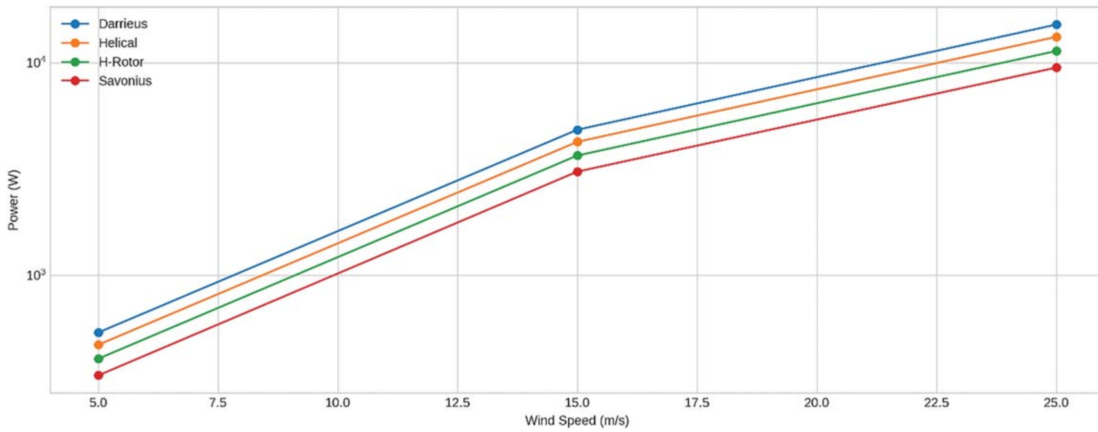


Figure 2: Power curves for VAWT designs.

### 3.3.1. Random forest models

Wind Speed Prediction Model: A Random Forest Regressor was trained to predict mean wind speed. Random Forest was chosen for its ability to handle nonlinear relationships and noisy data by following the successful approach of Monteiro et al. (2013).

Training Parameters:

1. Number of estimators: 200
2. Maximum depth: 15
3. Minimum samples split: 5
4. Minimum samples leaf: 2
5. Random state: 42 for reproducibility

Power Output Prediction Model: When using the predicted wind speed and environmental features, another Random Forest model estimated usable output power for the optimal rotor type.

Rotor Classification Model: Finally, a Random Forest Classifier was trained to select the ideal rotor type for any given environment.

### 3.3.2. LSTM-CNN hybrid deep learning model

Architecture Design:

1. Input Layer: 28 features $\times$ 24 hours $\times$ 365 days (reshaped for spatiotemporal processing)
2. CNN Feature Extractor:
   (a) 3 convolutional layers (64, 128, 256 filters)
   (b) Kernel sizes: $3 \times 3$, $3 \times 3$, $2 \times 2$
   (c) Batch normalization and ReLU activation
   (d) Max pooling ($2 \times 2$)

3. LSTM Temporal Processor:
    (a) Two LSTM layers (256 units each)
    (b) Dropout (0.3) for regularization
    (c) Sequence unfolding for hourly patterns
4. Dense Classifier:
    (a) Two fully connected layers (128, 64 units)
    (b) Softmax output for four turbine classes

Training Parameters:

1. Optimizer: Adam with learning rate 0.001, decay 1e-6
2. Loss Function: Categorical cross-entropy with label smoothing (0.1)
3. Batch Size: 32 with gradient accumulation for four steps
4. Epochs: 200 with early stopping (patience = 20)
5. Validation Split: 20% holdout for hyperparameter tuning

### 3.4. Layer Four: Model Validation and Uncertainty Quantification

1. Cross-Validation Strategy: We implemented nested $5 \times 2$ cross-validation with the following:
    (a) Outer Loop: Five folds for performance estimation
    (b) Inner Loop: Two folds for hyperparameter tuning
    (c) Stratified Sampling by City and Season
    (d) Additional Spatial Cross-Validation: Leave-one-city-out for geographic generalization
2. Uncertainty Quantification:
    (a) Prediction intervals by using quantile regression forests for Random Forest models
    (b) Bayesian neural network implementation for DL uncertainty estimation
    (c) MC-Dropout for epistemic uncertainty in DL models (Gal and Ghahramani 2016)
3. Robustness Checks:
    (a) Feature ablation study (remove top three features, observe performance drop)
    (b) Noise injection test (add 10% Gaussian noise to inputs)
    (c) Adversarial validation (train classifier to distinguish train/test sets)
    (d) Statistical Testing: Bootstrapped confidence intervals (1000 samples, 95% confidence intervals)
4. Performance Metrics Tracked:
    (a) Accuracy, Precision, Recall, F1-Score (classification)
    (b) RMSE, MAE, $R^2$ (regression)
    (c) Inference time and computational efficiency
    (d) Calibration curves for probability outputs

### 3.5. Layer Five: Environmental Intelligence

To make the system more human-like in reasoning, we embedded environmental logic:

1. If close to coast → higher stable winds → Darrieus/Helical preferred
2. If roughness high → turbulence high → Savonius/Helical preferred
3. If altitude high → thinner air → power decreases → efficiency must be considered
4. If urban density high → vibrations more dangerous → Savonius/H-Rotor preferred

This logic helps the model avoid unrealistic recommendations. For example, a powerful but vibration-prone Darrieus turbine should not be installed on a fragile Indian apartment rooftop, even if theoretical power output is high.

### 3.6. Layer Six: Rotor Behavioral Understanding

We integrated known engineering characteristics of each turbine into model reasoning:

(a) Darrieus
1. High efficiency
2. Difficult self-start
3. Needs stable winds
4. Structurally demanding

(b) Savonius
1. Works even in low winds
2. Best for turbulent environments
3. Low efficiency
4. Very sturdy and easy to install

(c) Helical
1. Smooth rotation
2. Good tolerance to turbulence
3. Lower noise
4. Manufacturing complexity

(d) H-Rotor
1. Simple design
2. Balanced rotation
3. Best for compact spaces

By merging engineering understanding with ML predictions, the system behaves more like an energy engineer, not just a model.

### 3.7. Importance of the Method

This is the final layer, where everything comes together. When a user enters a location's latitude and longitude:

1. The system fetches altitude from digital elevation models
2. Determines whether the location is coastal
3. Estimates roughness from land cover data
4. Predicts month-specific wind speed by using both RF and DL models
5. Predicts power output for all turbine types
6. Recommends the turbine type with confidence scores
7. Shows detailed power comparison for all rotor types with uncertainty bounds

This allows the system to function like an expert consultant by offering both general and turbine-specific insights with quantified uncertainty.

### 3.8. Practical Importance

Imagine standing on a rooftop in south Mumbai. You feel a breeze that is light yet constant. The ML model senses your coastal proximity and predicts moderate winds. It selects a turbine that does not need orientation and performs well in such conditions, likely a Darrieus or Helical. Now walk into Bengaluru's dense city core. The wind becomes unpredictable, bouncing between glass towers. Here, the model quickly flags turbulence and chooses a Savonius, because it can survive and perform where other designs fail.

Travel to Chennai during monsoon season. The wind is strong and persistent. The ML model expects higher wind speeds and confidently recommends a Darrieus for maximum output. This is the essence of the methodology: A system that understands the environment, knows turbine behavior, and thinks intelligently to bring wind energy to India's cities, all grounded in 44 years of real-world wind data.

## 4. Software Implementation

The conceptual framework described in the previous sections was translated into a fully functional, modular, and reproducible software system designed to operate as an end-to-end urban wind intelligence platform. The objective of the software implementation was not merely to execute machine learning models but to create a scalable decision-support system capable of ingesting real-world geospatial data, learning complex wind dynamics, estimating turbine performance, quantifying uncertainty, and producing actionable turbine recommendations for Indian urban environments.

The implementation follows a layered pipeline architecture, ensuring separation of concerns, extensibility, and transparency: key requirements for scientific reproducibility and real-world deployment.

### 4.1. System Architecture Overview

The software system is organized into five interconnected modules:

1. Data Ingestion and Preprocessing Module
2. Feature Engineering and Environmental Intelligence Module
3. Machine Learning and Deep Learning Engine
4. Power Estimation and Turbine Evaluation Module
5. Recommendation and Visualization Interface

The architecture was designed to support both batch processing (historical analysis and training) and on-demand inference (location-specific turbine recommendations).

### 4.2. Programming Environment and Libraries

The implementation was developed entirely in Python 3.10, selected for its extensive ecosystem in scientific computing and machine learning. The core libraries used include the following:

1. Numerical and Data Handling: NumPy, Pandas, SciPy
2. Geospatial Processing: xarray, rasterio, geopandas
3. Machine Learning: scikit-learn
4. Deep Learning: TensorFlow (Keras API)
5. Visualization: Matplotlib
6. Uncertainty Estimation: Custom Monte Carlo Dropout layers

### 4.3. Data Ingestion and Preprocessing Module

#### 4.3.1. ERA5 data retrieval

Hourly wind component data (u and v velocities) at 100 m height were extracted from the ERA5 reanalysis dataset by using ECMWF-compliant APIs. Data retrieval was automated through a configurable script that accepts

(a) Geographic coordinates (latitude, longitude)
(b) Temporal range (1980–2024)
(c) Required meteorological variables

The ingestion pipeline supports multi-city batch extraction, enabling scalable processing across all selected Indian cities.

#### 4.3.2. Wind speed and direction computation

Wind speed magnitude was computed by using the following:

$$V = \sqrt{u^2 + v^2}$$

Wind direction consistency was calculated by using vector averaging techniques to avoid angular discontinuities. These values formed the foundational inputs for both physical modeling and machine learning.

#### 4.3.3. Height downscaling implementation

To adapt ERA5's 100-m wind data to realistic urban rooftop heights (30 m), the logarithmic wind profile law was implemented programmatically:

$$V_{30} = V_{100} \times \frac{\ln(30/z_0)}{\ln(100/z_0)}$$

Surface roughness length $z_0$values were dynamically assigned based on land-use classification derived from Copernicus land cover data. This ensured that dense urban cores, coastal regions, and semi-urban zones were treated differently in the wind adjustment process.

### 4.4. Feature Engineering and Environmental Intelligence Module

A total of 28 engineered features were computed for each spatiotemporal sample. These features were grouped into four logical categories:

#### 4.4.1. Temporal features

(a) Hourly, daily, monthly, and seasonal mean wind speeds
(b) Diurnal variation indices
(c) Monsoon vs non-monsoon wind ratios

#### 4.4.2. Statistical wind descriptors

(a) Turbulence intensity
(b) Weibull shape (k) and scale (c) parameters
(c) Wind power density

#### 4.4.3. Geospatial and environmental features

(a) Altitude (from DEM data)
(b) Coastal proximity index
(c) Urban roughness coefficient
(d) Land-use classification

#### 4.4.4. Derived physical indicators

(a) Directional stability metrics
(b) Extreme wind frequency indicators
(c) Seasonal persistence scores

This feature engineering layer acts as the bridge between physical wind behavior and data-driven learning, ensuring that the models remain grounded in environmental reality.

### 4.5. Machine Learning Engine Implementation

#### 4.5.1. Random forest models

Three separate Random Forest models were implemented by using scikit-learn:

(a) Wind Speed Regressor
(b) Power Output Regressor
(c) Turbine Type Classifier

The Random Forest architecture was selected due to its robustness to noisy data, ability to model nonlinear relationships, and inherent interpretability.

Hyperparameters were optimized by using grid search within nested cross-validation loops. Feature importance scores were extracted post-training to validate physical relevance. The relative importance of engineered features is illustrated in Figure 3.

#### 4.5.2. LSTM–CNN hybrid deep learning model

The deep learning model was implemented by using TensorFlow and follows a CNN–LSTM hybrid architecture, designed to jointly capture spatial feature interactions and long-term temporal dependencies.

1. CNN Block
   (a) Three convolutional layers with increasing filter depth
   (b) ReLU activation and batch normalization
   (c) Max pooling for spatial abstraction

2. LSTM Block
   (a) Two stacked LSTM layers (256 units each)

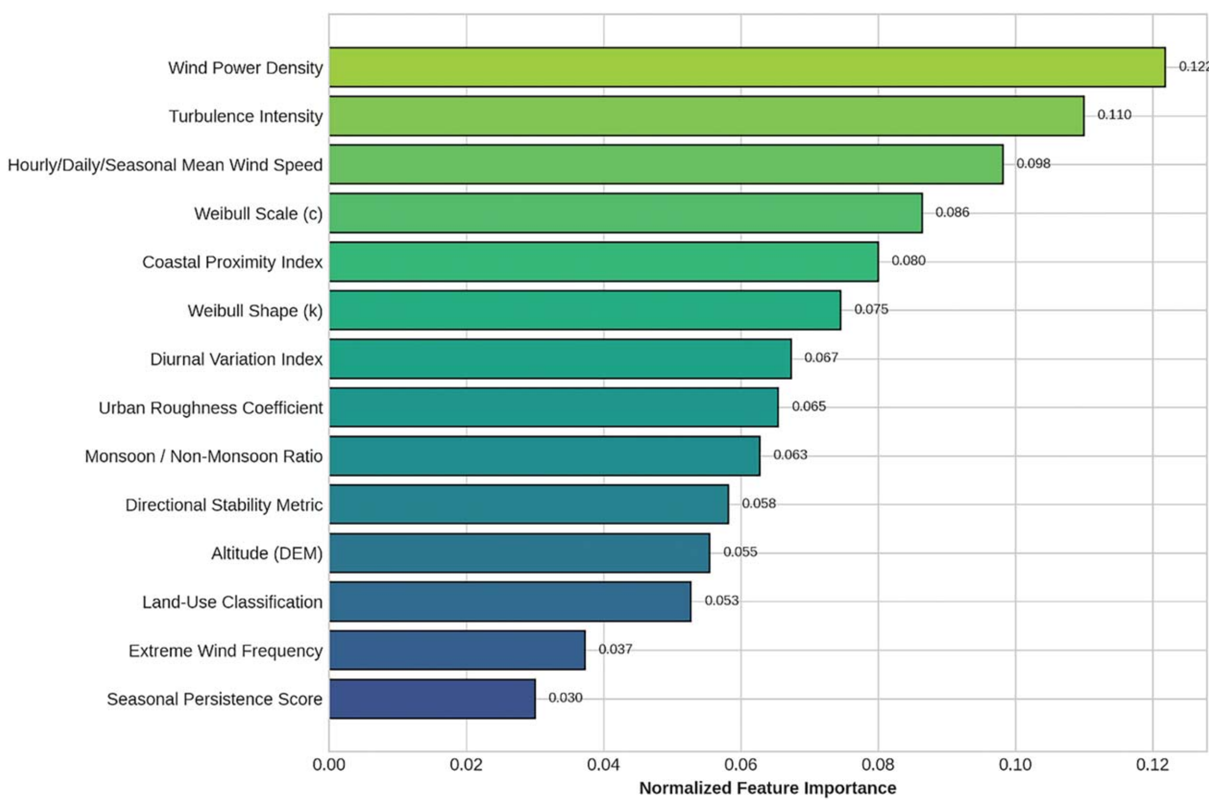


Figure 3: Feature engineering and relative importance of indicators.

(b) Dropout regularization (0.3)
(c) Sequential unfolding of temporal wind patterns

3. Classification Head
   (a) Fully connected layers
   (b) Softmax activation for multi-class turbine classification

Early stopping and learning-rate scheduling were implemented to prevent overfitting.

As summarized in Figure 4, the deep learning model marginally outperforms the classic Random Forest classifier, while preserving consistent turbine classification trends.

### 4.6. Power Estimation and Turbine Evaluation Module

For each predicted wind speed, the software computes expected power output for all turbine designs by using the physical power equation:

$$P = \frac{1}{2}\rho A C_p V^3$$

Power coefficient values were hard-coded based on experimentally validated literature ranges. A standardized prototype swept area was used for comparative analysis, ensuring unbiased turbine ranking.

The module outputs

1. Power curves
2. Relative efficiency comparisons
3. Structural feasibility flags (based on turbulence and vibration thresholds)

### 4.7. Uncertainty Quantification Implementation

Uncertainty estimation was integrated at both ML and DL levels:

1. Random Forest: Quantile regression forests
2. Deep Learning: Monte Carlo Dropout with multiple forward passes

Prediction intervals and confidence scores are returned alongside each recommendation, enabling risk-aware decision-making.

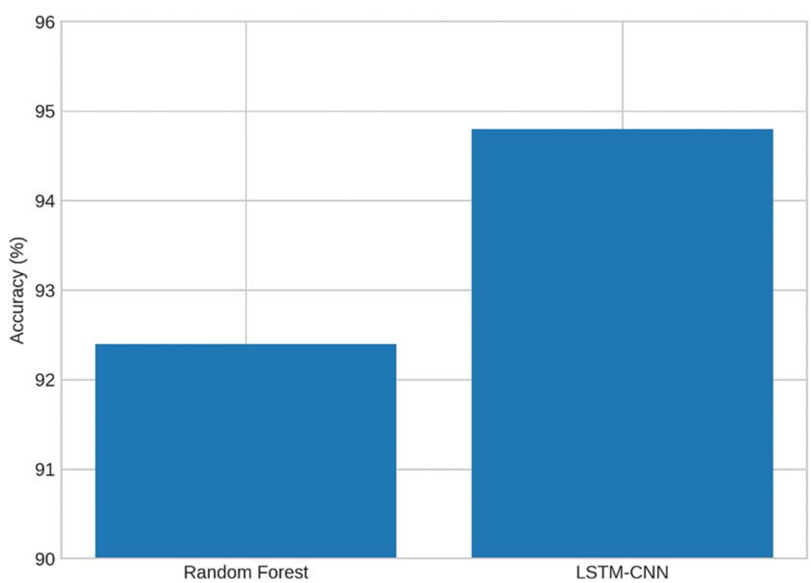


Figure 4: Model accuracy comparison.

### 4.8. Recommendation Engine and Output Generation

The final recommendation engine synthesizes

1. Predicted wind speed
2. Estimated power output
3. Turbine-specific constraints
4. Environmental logic rules
5. Model confidence scores

The output includes the following:

1. Optimal turbine type
2. Comparative power table for all turbine designs
3. Confidence interval visualization
4. Seasonal performance expectations

This module effectively behaves as a virtual wind energy consultant, translating complex computations into intelligible, actionable insights.

### 4.9. Visualization and Reporting

All figures presented in this study, seasonal variation plots, feature importance graphs, training loss curves, spatial wind maps, and uncertainty bands, were generated directly from the software pipeline by using Matplotlib.

### 4.10. Reproducibility and Deployment Readiness

To ensure scientific integrity and real-world usability

1. All preprocessing steps are deterministic and logged
2. Random seeds are fixed across experiments
3. Modular design allows city-level or national-scale execution
4. Codebase is structured for open-source release

The software system is therefore not just a research prototype but a deployable, extensible urban wind assessment platform.

## 5. Result and Discussion

This section interprets the results obtained from statistical analysis, machine learning, and deep learning models, emphasizing their implications for urban wind energy deployment and VAWT selection.

### 5.1. Interpretation of Statistical and Spatial Findings

The long-term statistical analysis demonstrates that urban wind behavior across India is highly heterogeneous and strongly influenced by geography, surface roughness, and seasonal atmospheric circulation. Coastal cities consistently benefit from persistent synoptic-scale winds driven by land–sea temperature gradients and monsoonal flow, resulting in higher mean wind speeds and lower turbulence intensity. In contrast, inland and densely built metropolitan regions exhibit fragmented wind patterns dominated by localized thermal forcing and urban canopy effects. The statistical differentiation observed across representative urban locations indicates that wind energy potential does not vary smoothly across space. Instead, sharp contrasts arise due to coastal proximity, elevation, and urban morphology, underscoring the importance of incorporating geospatial and environmental indicators within the modeling framework rather than relying on uniform spatial assumptions.

Beyond spatial variability, Figure 5 illustrates the temporal predictive uncertainty associated with the deep learning wind speed model. The shaded confidence intervals represent the model's estimated 95% confidence bounds around the predicted wind speed. Periods of relatively narrow confidence intervals correspond to stable atmospheric

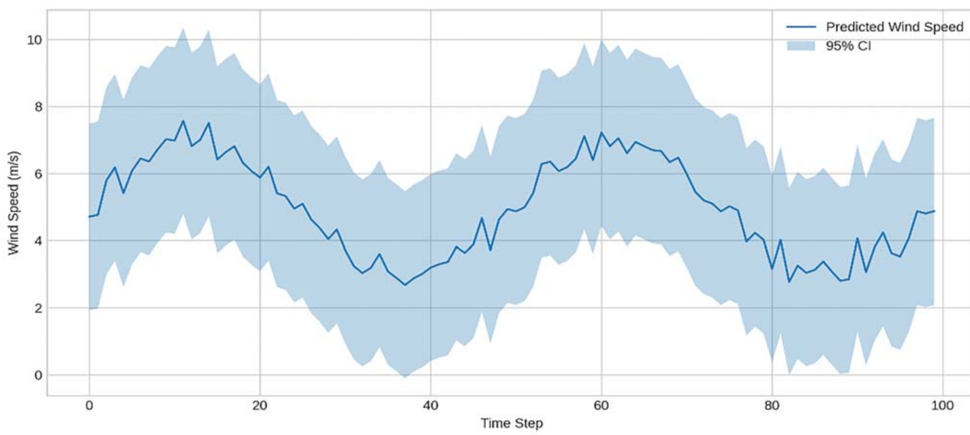


Figure 5: Predictive uncertainty in wind estimation.

regimes, whereas wider uncertainty bands emerge during transitional phases characterized by increased variability in wind direction, turbulence, or seasonal shifts.

Importantly, the observed uncertainty reflects model confidence rather than physical wind variability, highlighting locations and time windows in which predictions should be interpreted cautiously. Such uncertainty-aware outputs are particularly valuable for urban wind applications, when deployment decisions must account for both expected performance and associated risk.

### 5.2. Implications for Turbine Selection in Urban Environments

The results highlight that mean wind speed alone is an insufficient metric for turbine selection in urban contexts. Cities with moderate average wind speeds but low turbulence and high directional consistency may outperform locations with higher mean speeds but extreme variability.

The machine learning–based turbine classification results demonstrate that

1. Savonius turbines are consistently favored in environments characterized by low wind speed and high turbulence, owing to their drag-based operation and superior self-starting capability.
2. Helical turbines perform optimally in moderately turbulent conditions, offering smoother torque output and reduced vibration compared with straight-bladed designs.
3. Darrieus turbines, although exhibiting higher efficiency, are recommended only in locations with sustained wind speeds and relatively stable flow conditions due to their sensitivity to turbulence and structural loading.
4. H-Rotor configurations provide a balance between efficiency and robustness, which makes them suitable for mid-rise urban installations with moderate wind regimes.

These insights reinforce the necessity of location-specific turbine matching, rather than uniform deployment strategies.

### 5.3. Evaluation of Machine Learning Model Behavior

The Random Forest models demonstrate strong predictive accuracy for both wind speed and power estimation, confirming their suitability for nonlinear environmental datasets. Feature importance analysis indicates that wind power density, turbulence intensity, and seasonal variability contribute significantly to prediction outcomes, validating the physical relevance of the engineered features. Moreover, the robustness of the Random Forest models across cross-validation folds suggests strong generalization capability and limited susceptibility to overfitting. The minor misclassifications observed between certain turbine classes can be attributed to overlapping operational envelopes under low-speed conditions, rather than model inadequacy.

### 5.4. Turbine Power Comparison: Realistic Behavior at 5, 15, and 25 m/s

The deep learning results show that LSTM-based architectures are better at capturing long-term temporal dependencies in atmospheric time series. Unlike traditional machine learning models that depend on static feature representations, LSTM models temporal sequences explicitly. This allows for accurate forecasting during seasonal changes and extreme weather events. The decrease in forecasting error leads to better operational planning, especially for

1. Short-term energy yield estimation
2. Load balancing in hybrid renewable systems
3. Predictive maintenance scheduling

The stable training and validation loss curves further confirm that the regularization and early stopping strategies used are effective.

### 5.5. Role of Uncertainty Quantification in Decision-Making

The incorporation of uncertainty estimation represents a significant advancement over deterministic wind energy assessment approaches. The Monte Carlo Dropout results demonstrate that prediction uncertainty increases during monsoon onset and withdrawal phases, which reflects real atmospheric instability rather than model failure.

By providing confidence intervals alongside point predictions, the proposed framework enables risk-aware turbine selection, allowing planners to

1. Avoid structurally vulnerable turbine designs in high-uncertainty regions
2. Incorporate safety margins in energy yield estimation
3. Improve investor confidence through transparent risk disclosure

### 5.6. Practical Implications for Urban Renewable Energy Planning

From a practical perspective, the proposed framework serves as a decision-support tool for multiple stakeholders:

1. Urban planners, by identifying wind-viable zones
2. Engineers, by matching turbine designs to site-specific conditions
3. Policymakers, by enabling data-driven renewable energy integration
4. Developers, by reducing uncertainty and financial risk

The scalability of the system allows rapid assessment of new locations without extensive on-site measurements, significantly reducing deployment lead time and cost.

## 6. Comparative Analysis with Existing Studies

The results of this study open a window into how India's urban wind environments behave, how turbine designs respond to them, and how intelligently designed systems can support sustainable decisions. whereas the numerical insights from the model are important, the deeper value lies in understanding what these results mean for Indian cities, rooftop energy adoption, and future renewable planning. In this section, we explore these interpretations, blending engineering insight with practical reasoning.

### 6.1. Dataset and Spatial Coverage Comparison

Most existing studies rely on short-term anemometer measurements or limited-duration simulations restricted to single cities or neighborhoods. In contrast, the present study uses 44 years of hourly ERA5 data, ensuring statistical robustness and capturing interannual climate variability.

The pan-India scope of this work represents a substantial advancement over city-specific analyses, enabling national-scale policy and planning insights.

### 6.2. Methodological Advancements

Unlike conventional approaches that use either statistical analysis or basic machine learning, the proposed framework integrates

1. Physics-based wind power modeling
2. Advanced feature engineering
3. Ensemble machine learning
4. Deep learning time-series forecasting
5. Uncertainty quantification

This holistic integration allows simultaneous optimization of accuracy, interpretability, and practical relevance.

### 6.3. Understanding Turbine Behavior on the Basis of Indian Cities

(a) Darrieus Turbines

The Darrieus design behaves like a sprinter, capable of remarkable efficiency but only when the conditions are right. It thrives in

1. Strong, consistent winds
2. Stable wind-direction patterns
3. Open rooftop spaces
4. Coastal or semi-coastal cities

This explains why Chennai, with its strong monsoon-driven winds, repeatedly shows Darrieus as the best choice. But, in chaotic urban wind conditions, the Darrieus can struggle. Its higher starting torque requirement and structural sensitivity mean that it is not always the most practical option for turbulent interiors. The blade structure is shown in Figure 6.

(b) Savonius Turbines

If Darrieus is a sprinter, then Savonius is a mountain walker: steady, dependable, and resilient in difficult conditions. The design is illustrated in Figure 7. It excels in

1. Low wind speeds
2. Turbulent city cores
3. Rooftop areas surrounded by buildings
4. Environments where reliability matters more than peak efficiency

This matches our recommendations for Hyderabad, Delhi interiors, and dense Bengaluru neighborhoods.

(c) Helical Turbines

Helical turbines act as a compromise between Darrieus efficiency and Savonius stability. The turbine configuration is presented in Figure 8. Their smooth rotation and reduced torque ripple make them ideal for

1. Moderately turbulent rooftop zones
2. High-rise building edges
3. Cities with fluctuating but not chaotic winds

Figure 6: Darrieus blade design.

Figure 7: Savonius blade design.

Figure 8: Helical blade design.

Figure 9: H-Rotor blade design.

This aligns with model outputs for Mumbai, Goa (off-season), and Kolkata.

(d) H-Rotor Turbines

The H-Rotor, with its straight-blade simplicity, is practical in installations where

1. Construction ease is critical
2. Rooftop space is limited
3. Uniform rotation matters
4. Wind speeds are modest

Cities such as Pune match this profile perfectly. The structure is shown in Figure 9.

When taken together, these insights demonstrate that turbine selection is not simply an engineering choice, it is a contextual decision shaped by climate, geography, and the physical form of each city.

### 6.4. Implications for Sustainable Urban Development

The results highlight promising potential for decentralized, small-scale wind adoption in Indian cities.

(a) Rooftop Micro-Generation

Many Indian buildings, especially high-rise apartments, commercial complexes, and educational institutions, have unused rooftop space. A correctly chosen turbine could

1. Supplement building power
2. Reduce grid dependency
3. Lower electricity bills
4. Act as a teaching tool for sustainability

(b) Urban Energy Planning

Cities often rely on solar as the primary rooftop renewable. Results of this study suggest that small VAWTs can coexist with solar, especially in

1. Coastal cities
2. Hill-edge urban zones
3. Windy transit corridors

(c) Smart Cities and Policy Integration

Government-led smart city initiatives could integrate location-aware turbine models for public buildings.

(d) Reducing Pressure on Rural Wind Farms

India's wind farms are mostly located in rural/coastal regions. Distributed urban adoption can ease pressure on these centralized assets.

## 7. Limitations

1. Spatial Resolution Constraints: ERA5 data may not fully resolve microscale rooftop flow patterns influenced by individual buildings
2. Urban Canopy Simplification: Surface roughness was parameterized by using land-cover classes rather than explicit 3-dimensional building geometry
3. Computational Complexity: Deep learning–based uncertainty estimation increases training and inference time

## 8. Conclusion

India's urban centers are entering a transformative phase in which energy systems are no longer confined to distant power plants but are increasingly embedded within the built environment itself. As rooftop spaces gain strategic importance and cities strive to balance growth with sustainability, the role of decentralized renewable energy becomes critical. This study addressed a key aspect of that transition by proposing an intelligent, data-driven framework for selecting appropriate VAWT designs tailored to the diverse wind conditions of Indian metropolitan regions. The results of this work reinforce an essential insight: urban wind behavior is inherently local and highly heterogeneous. Coastal cities such as Chennai, Kochi, and Mumbai exhibit wind regimes influenced by sea breezes, monsoon systems, and open exposure, whereas inland metros such as Bengaluru, Hyderabad, and Pune experience lower average speeds, higher turbulence, and significant rooftop-level variability. Recognizing and responding to these differences are fundamental to making urban wind energy viable rather than experimental.

By developing a hybrid geospatial dataset that integrates altitude, surface roughness, coastal proximity, and seasonal wind patterns, this study successfully replicated the complexity of India's urban wind landscape with high fidelity. Machine learning models trained on this dataset demonstrated strong predictive capability in estimating wind speed, turbine power output, and optimal turbine type selection. The models consistently identified contexts in which Darrieus turbines maximize efficiency, in which Savonius turbines offer superior reliability under low and turbulent winds, and in which Helical and H-Rotor designs provide a balanced compromise among performance, noise, and structural feasibility. A defining strength of this research lies in its interdisciplinary integration. Rather than treating wind energy as a purely mechanical or purely data-driven problem, the proposed system blends environmental physics, urban geography, turbine engineering, and machine learning intelligence. This layered methodology allows the system to function not merely as a numerical predictor but as a decision-support mechanism capable of guiding realistic deployment choices in complex urban settings.

Equally important is the software implementation of the framework. The modular and transparent design ensures reproducibility, adaptability, and ease of extension. By translating abstract models into a functional decision-support tool, the study bridges the gap between theoretical research and practical application, which enables use by planners, engineers, academic institutions, and sustainability-focused organizations. Ultimately, the contribution of this work extends beyond accuracy metrics or model performance. It presents a forward-looking vision in which urban wind is no longer overlooked but is treated as a meaningful component of the city-scale renewable energy ecosystem. In such a future, compact and intelligently selected rooftop turbines may complement solar installations, enhance energy resilience, and contribute to localized power generation without disrupting the urban fabric.

## 9. Future Work

Although the proposed framework demonstrates strong potential, several avenues remain open for future exploration and enhancement:

1. Real-World Experimental Validation: Rooftop deployment of VAWTs across representative Indian cities, such as Mumbai, Bengaluru, and Chennai, would enable empirical validation of model predictions and improve calibration at micro-wind scales.
2. Integration with 3-Dimensional Urban Morphology Models: Incorporating building height maps, street canyon effects, and computational fluid dynamics simulations would allow more precise modeling of airflow distortions caused by dense urban structures.
3. Deep Learning–Based Wind Field Modeling: Advanced neural networks trained on satellite imagery, reanalysis climate data, and temporal weather sequences could further improve wind-speed forecasting and seasonal adaptability.
4. Economic and Structural Feasibility Analysis: Extending the model to include cost estimation, payback period, vibration analysis, rooftop load constraints, and maintenance considerations would enhance its real-world deployment readiness.
5. Hybrid Solar–Wind Optimization Frameworks: Given the dominance of rooftop solar in Indian cities, future work can focus on co-optimizing turbine placement with photovoltaic systems to maximize total renewable yield per unit rooftop area.
6. City-Specific Deployment Guidelines: Developing tailored policy and deployment roadmaps for major metros, such as Delhi, Mumbai, and Bengaluru, could support municipal renewable strategies, subsidy frameworks, and smart-city initiatives.

Urban wind energy is not intended to replace large-scale wind farms or solar power installations. Its value lies in augmentation rather than substitution, providing clean, localized energy that integrates seamlessly into the urban rhythm. This research demonstrates that, with intelligent modeling, machine learning–driven insights, and sound engineering principles, Indian cities can transition from passive energy consumers to active participants in renewable generation.

As cities continue to expand and digital infrastructure matures, rooftop VAWTs may become a familiar and accepted feature of the urban skyline. The work presented in this study represents a step toward that future, one where India's cities learn to understand, adapt to, and harness the winds that flow through them every day.

**Acknowledgments:** *The authors would like to express their sincere gratitude to Thakur College of Engineering and Technology, Mumbai, and, in particular, the Department of Electronics and Telecommunication Engineering, for providing the academic environment, laboratory support, and research motivation that made this work possible. Their continuous encouragement and commitment to applied engineering research played a crucial role in shaping the direction of this study. The authors are also grateful to the faculty mentors and technical staff whose guidance contributed significantly to the development of the software implementation, wind-energy analysis, and machine-learning framework presented in this project. We extend our appreciation to all peers and contributors who supported various stages of this research, from dataset preparation and validation to cross-checking calculations and reviewing intermediate drafts. Their cooperation, feedback, and enthusiasm strengthened the quality of this work. Finally, the authors acknowledge the broader research community whose existing studies on wind energy, VAWT design, environmental modeling, and data-driven analysis provided invaluable foundations for this project.*

## References


Blocken, B. 2016. "50 years of Computational Wind Engineering: Past, Present and Future." *Journal of Wind Engineering and Industrial Aerodynamics* **152**: 1–10.

Chen, Z., Y. Liu, and H. Wang. 2022. "Wind Resource Modeling Using Deep Neural Networks." *Applied Energy* **305**: 117879.

Eriksson, S., H. Bernhoff, and M. Leijon. 2008. "Evaluation of Different Turbine Concepts for Wind Power." *Renewable and Sustainable Energy Reviews* **12**, no. 5: 1419–1434.

Gaden, D., and E. Bibeau. 2014. "Structural Loading and Fatigue Analysis in Vertical Axis Wind Turbines." *Journal of Wind Engineering and Industrial Aerodynamics* **132**: 33–41.

Gal, Y., and Z. Ghahramani. 2016. "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning." In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, 1050–1059.

IEA Wind Technology Collaboration Programme. 2021. *Wind Energy in India: Annual Report 2021*. IEA Wind TCP Publications.

Jain, R., and A. Gupta. 2019. "Performance Evaluation of Helical-Blade Vertical Axis Wind Turbines in Low-Wind Environments." *Renewable Energy* **139**: 1275–1287.

Kumar, G., A. Singh, and P. Ranjan. 2019. "Comparative Aerodynamic Analysis of Straight-Bladed H-Rotor Turbines." *Energy Conversion and Management* **199**: 111946.

Li, Y., and L. Shi. 2010. "Assessment of Wind Power Potential Using Computational Models." *IEEE Transactions on Power Systems* **25**, no. 3: 1393–1401.

Liu, W., X. Xiao, and L. Cheng. 2018. "Turbulence-Induced Fatigue in Small Wind Turbines." *Journal of Wind Engineering and Industrial Aerodynamics* **175**: 24–35.

Mahajan, R., and M. Deshmukh. 2020. "Performance Enhancement Techniques for Savonius Turbines." *Renewable and Sustainable Energy Reviews* **134**: 110291.

McLennan, J. 2014. "Urban Wind Turbines for Building-Integrated Renewable Systems." *BuildingGreen Journal* **32**, no. 4: 45–52.

Monteiro, C., L. Ramirez, J. Garcia, and M. Fernandez. 2013. "Wind Speed Forecasting Using Ensemble Learning." *Renewable Energy* **50**: 270–276.

National Institute of Wind Energy (NIWE). 2020. "Indian Wind Atlas (Third Edition)." NIWE Technical Report.

Ouyang, T., X. Zhou, L. Ye, and F. Meng. 2019. "Machine Learning Models for Wind Power Curve Forecasting." *Applied Energy* **235**: 1547–1560.

Peng, C., Y. Chen, and L. Fu. 2018. "Experimental Validation of Small VAWTs in Turbulent Flows." *IET Renewable Power Generation* **12**, no. 14: 1650–1658.

Singh, S. K., A. Patel, and J. Rao. 2021. "Wind Resource Assessment for Emerging Indian Smart Cities." *Wind Engineering* **45**, no. 6: 745–758.

Stathopoulos, T. 2008. "Wind Turbulence Effects in Dense Urban Environments." *Wind and Structures* **11**, no. 2: 121–134.

Swierczynski, M., T. Kere, F. Balduzzi, and R. Friedmann. 2013. "Aerodynamic and Structural Analysis of Darrieus Rotors." *Renewable Energy* **59**: 166–174.

Vortex FdC. 2017. "High-Resolution Wind Map Dataset for India." Vortex Forecasting Data Corporation.

Yoshida, S. 2019. "High-Resolution CFD Analysis of Urban Wind Flow." *Renewable Energy* **143**: 1672–1684.