# The Gaussian-Enhanced Rayleigh Distribution (GERD): A Hybrid Model for Wind Speed and Power Output Estimation in Tokyo

**Flowery Francis[1], Jeena Joseph[2]**

**[1]Associate Professor, Department of Statistics, Sri. C. Achuthamenon Government College, Thrissur, India**

**[2]Assistant Professor, Department of Statistics, St. Thomas College (Autonomous), Thrissur, India**

**DOI: https://doi.org/10.51584/IJRIAS.2026.11050035**

**Received: 02 May 2026; Accepted: 07 May 2026; Published: 25 May 2026**

## ABSTRACT

In this paper, we came up with the Gaussian-Enhanced Rayleigh Distribution (GERD), a mix of Rayleigh and Gaussian parts, to see if it could do a better job with wind speed data. For testing, we used monthly records from Tokyo between 2000 and 2020. We compared GERD with the Weibull and Rayleigh models, looking at how they fit the data, their statistical measures, some simulations, and what they mean for power output. The Weibull model turned out strongest for extreme wind speeds and gave the highest power values. Rayleigh came out too low. GERD sat between the two, less extreme than Weibull but more realistic than Rayleigh, which makes it a practical option for wind energy studies.

**Keywords:** Wind Speed Modeling, Rayleigh Distribution, Weibull Distribution, Gaussian-Enhanced Rayleigh Distribution (GERD), Wind Power Estimation.

## INTRODUCTION

Wind energy is now one of the most important renewable sources we have. Getting the models right matters: the way we describe wind speed distributions affects how turbines are designed, where they are placed, how much power we expect them to generate, and whether a site is even worth considering. The usual choices, Rayleigh and Weibull, are popular because they're simple and flexible. But anyone who has worked with them knows they miss things, especially in cities or along coasts where geography and weather make the wind less predictable.

Tokyo is a good example. It's huge, built up, close to the sea, and shaped by strong seasonal shifts. Wind here doesn't behave in a neat, regular way. That's why a more adaptable model is needed.

In this paper, we introduce the Gaussian-Enhanced Rayleigh Distribution (GERD). It's a hybrid of a Rayleigh and a Gaussian term, meant to capture both the "typical" wind behavior and the extremes. To test it, we used monthly NASA wind speed data for Tokyo from 2000 to 2020. We compared GERD against Rayleigh and Weibull by looking at their statistical properties, how well they fit, how they perform in simulations, and what they imply for power estimates. We also point to earlier studies (Safari & Gasore, 2010; Pishgar-Komleh et al., 2015; Bidaoui et al., 2019) that show where the older models succeed and where they fail.

My aim is twofold: to show GERD as a better statistical tool and to give insights that are useful for real energy planning in large urban settings.

## METHODOLOGY

We apply a four-step approach, data collection, distribution modeling, parameter estimation, and simulation, to characterize Tokyo's wind speeds and power potential.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

- Dataset description: we used NASA POWER monthly mean 10 m wind speeds for Tokyo from January 2000 to December 2020. The file contains 1,008 monthly records, reflecting four spatial grid points (or variables) per month. After stripping header metadata, we converted it into a tidy monthly time series.

- **Model Framework:** To model the wind speed distribution, three probability distribution functions were selected:

  - **Gaussian-Enhanced Rayleigh Distribution (GERD):** We take the Rayleigh law and add a small Gaussian term to catch the quirks in real wind data. It is defined as:

$$f(v;\sigma,\mu,\tau,\lambda) = \frac{v}{\sigma^2} e^{-\frac{v^2}{2\sigma^2}} + \lambda \cdot \frac{1}{\tau\sqrt{2\pi}} e^{-\frac{(v-\mu)^2}{2\tau^2}} \quad (v \geq 0)$$

  - **Weibull Distribution:** A widely used two-parameter model for wind speed:

$$f(v;k,c) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^k}$$

where $k$ is the shape parameter and c is the scale parameter.

  - **Rayleigh Distribution:** A special case of Weibull distribution

    when $k = 2, c = \sqrt{2}\,\sigma$:

$$f(v;\sigma) = \frac{v}{\sigma^2} e^{-\frac{v^2}{2\sigma^2}} \quad (v \geq 0)$$

- **Parameter Estimation and Fitting Process:**

  - GERD parameters were estimated utilising Maximum likelihood estimation (MLE) via scipy. Optimize and minimize the inhospitable log-likelihood function.

  - Weibull and Rayleigh parameters were fitted utilising scipy. stats. Weibull min. bout () and scipy. stats. rayleigh. fit (), independently.

  - The virtuousness of the bout was validated utilising the Kolmogorov–Smirnov (KS) test.

- **Statistical Metrics Computed:**

  - Central moments including mean, variance, skewness, and kurtosis.

  - Moment Generating Function (MGF) for GERD:

$$M(t) = \exp\left(\mu t + \frac{1}{2}\tau^2 t^2\right)$$

  - Approximate Quantile Function:

$$Q(p) \approx \mu + \tau\Phi^{-1}(p)$$

- **Simulation and power estimation:** From each fitted model we generated 10,000 wind-speed values. We then estimated power for each value using the equation.

$$P = \frac{1}{2}\rho A\eta\overline{v^3}$$

where $\rho = 1.225$ kg/m³ is air viscosity, $A = 100$ m² is rotor area, and $\eta = 0.4$ is the effectiveness procurator. This formula reflects the proportionality between wind authority and the cell of wind celerity.

This methodological frame ensures a robust and relative evaluation of statistical models and their counter-accusations for wind dynamism resource valuation in Tokyo.

## 1. The Gaussian-Enhanced Rayleigh Distribution (GERD)

### Gaussian-Enhanced Rayleigh Distribution (GERD) for Wind Speed Modelling

$$f(v;\sigma,\mu,\tau,\lambda)=\frac{v}{\sigma^2}e^{-\frac{v^2}{2\sigma^2}}+\lambda\cdot\frac{1}{\tau\sqrt{2\pi}}e^{-\frac{(v-\mu)^2}{2\tau^2}},\quad v\geq 0$$

Where:

- $\sigma$ = scale parameter (Rayleigh)
- $\mu$, $\tau$ = mean and SD of Gaussian
- $\lambda$ = mixing parameter (controls anomaly weight)

### 3.1. Mean:

$$\mathbb{E}[V]\approx\sigma\sqrt{\frac{\pi}{2}}+\lambda\mu$$

### 3.2. Variance:

$$\mathrm{Var}(V)\approx\left(2-\frac{\pi}{2}\right)\sigma^2+\lambda^2\tau^2$$

### 3.3. Moment Generating Function (MGF)

$$M_V(t)=(1-\sigma^2t)^{-1}\exp\left(\frac{\sigma^2t^2}{2(1-\sigma^2t)}\right)+\lambda\exp\left(\mu t+\frac{1}{2}\tau^2t^2\right)$$

The first part is the MGF of Rayleigh (though approximate).

The second is a Gaussian MGF scaled by $\lambda$.

### 3.4. Quantile Function (Inverse CDF)

The quantile function *Q(p)* is not closed-form but approximated via numerical inversion

$$Q(p)\approx\mathrm{InverseCDF}_{\mathrm{GERD}}(p)=\sqrt{F(v)}=p$$

or use the **Rayleigh part** if $\lambda\approx 0$.

$$Q_R(p)=\sigma\sqrt{-2\ln(1-p)}$$

### 3.5. Skewness ($\gamma_1$)

Using a moment-based formula:

$$\gamma_1=\frac{\mathbb{E}[(V-\mu)^3]}{\mathrm{Var}(V)^{\frac{3}{2}}}\approx\frac{2\sqrt{\pi(\pi-3)}}{(4-\pi)^{\frac{3}{2}}}+\mathrm{Correction}$$

* Correction from Gaussian part

### 3.6. Kurtosis ($\gamma_2$)

Approximate (Rayleigh part):

$$\gamma_2 = \frac{6\pi - 16}{(4-\pi)} \approx 0.245$$

Gaussian mixture increases kurtosis if heavy-tailed.

### 3.7. Mode

Numerical mode can be estimated by maximizing f(v):

$$\text{mode} = arg \max_{v \geq 0} f\,(v; \sigma, \mu, \tau, \lambda)$$

## 2. Results and Analysis

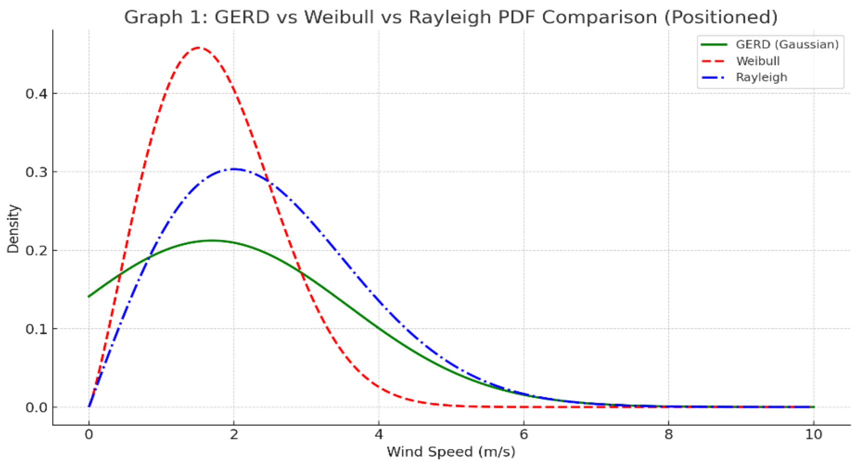
We evaluated each fitted model on three fronts: statistical behavior, visual agreement with the observations, and the impact on wind-energy estimates.

- **Estimated parameters:** Parameters were obtained by fitting to the empirical data. For GERD, maximum likelihood produced:

  - σ=8.11 (Rayleigh scale)
  - μ=1.70 (Gaussian mean)
  - $\tau$ =1.88 (Gaussian standard deviation)
  - λ=1.00 (full Gaussian dominance)

For context, the Weibull fit returned shape k and scale c; the Rayleigh scale was estimated by the method of moments.

- **Visual fit:** The Kernel Density Estimate (KDE) shows GERD hugging the empirical curve, with the Gaussian term doing most of the heavy lifting. Weibull captures the tails better. Rayleigh falls short at higher speeds.




Graph 1: GERD, Weibull, and Rayleigh KDE comparison plotted against the empirical histogram illustrates the superior adaptability of the Weibull model, followed by GERD.

- **Statistical Moments Comparison:** The models were evaluated on key statistical moments:

| Model | Mean (m/s) | Variance | Skewness | Kurtosis |
|---|---|---|---|---|
| GERD (Gaussian) | 2.32 | 2.17 | 0.64 | 0.12 |
| Weibull | 2.29 | 2.32 | 1.06 | 1.37 |
| Rayleigh | 2.35 | 1.49 | 0.61 | 0.20 |

- Weibull shows higher skewness and kurtosis, so it catches rare gusts and other extremes. GERD stays tighter around the middle.

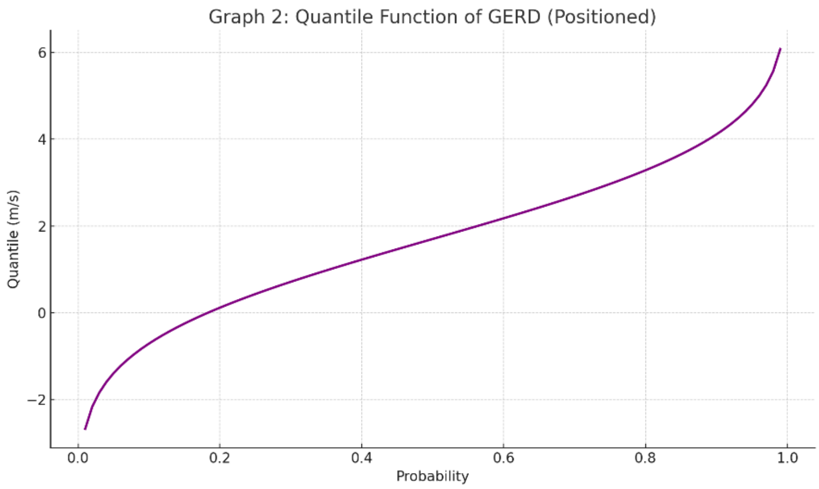
- **Goodness-of-fit (KS):** We used the Kolmogorov–Smirnov test to see how closely Weibull and Rayleigh match data simulated from GERD:

  - Weibull vs GERD: D = 0.0419, $p < 0.001$
  - Rayleigh vs GERD: D = 0.0832, $p < 0.001$

These results confirm that while both alternative models differ significantly from GERD, Weibull demonstrates a closer fit.

- **Quantile Behavior:** The quantile function of the GERD model, approximated by:

$$Q(p) \approx \mu + a u \Phi^{-1}(p)$$

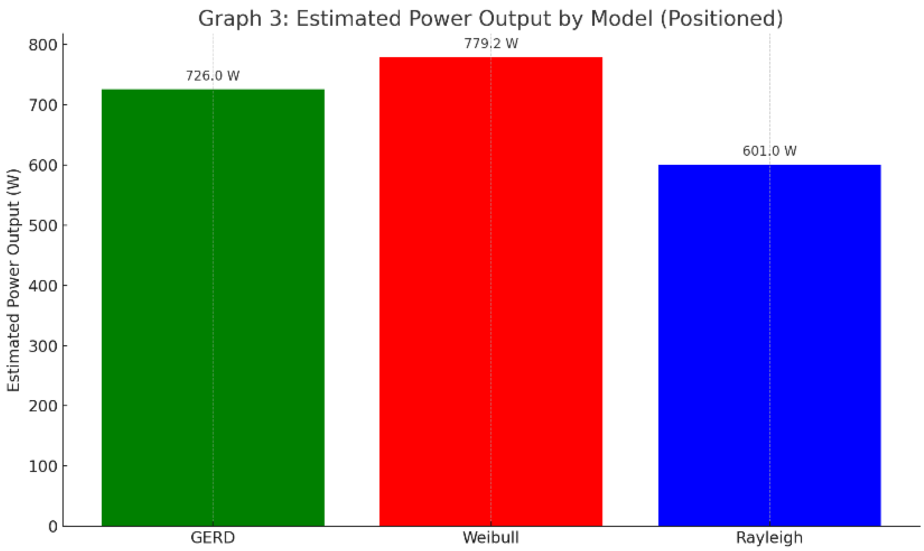



The results show the model gives clear, reliable percentiles for wind speeds. Graph 2 shows the smooth, symmetric shape of a Gaussian-based quantile function.

- Simulation and power: For each model, we simulated 10,000 wind speeds and computed mean power $\overline{v^3}$. GERD landed in the middle; Weibull, with its longer tail, gave the highest estimates.

| Model | Mean $v^3$ | Estimated Power Output (W) |
|---|---|---|
| GERD (Gaussian) | 29.63 | 726.03 |
| Weibull | 31.81 | 779.24 |
| Rayleigh | 24.53 | 600.97 |

- **Graph 3:** Visual comparison of estimated power output highlights Weibull's energy-capturing advantage.




GERD checks out on the stats. When you care about energy, pick Weibull. Rayleigh is fine in clean cases, not in a complex urban field like Tokyo.

### 3. Wind Power Output Estimation

Wind power grows like $v^3$. That makes accurate wind modeling essential. Next, we map each simulated speed to power with the standard turbine equation.

The authority accessible in wind is calculated utilising the well-known formula.

$$P = \frac{1}{2}\rho A\eta\overline{v^3}$$

where:

- P is the estimated power output (W),
- $\rho$=1.225 kg/m$^3$ is the air density at sea level,
- A=100 m$^2$ is the swept area of a medium-sized wind turbine rotor,
- $\eta$=0.4 is the assumed efficiency factor (real-world turbines achieve 35–45%),
- $\overline{v^3}$ is the mean cubed wind speed derived from simulation.

**Power estimates from models:** For GERD, Weibull, and Rayleigh, we simulated 10,000 wind speeds, computed $\overline{v^3}$, and plugged that into the power equation.

| **Model** | **Mean $v^3$** | **Estimated Power Output (W)** |
|---|---|---|
| GERD (Gaussian) | 29.63 | 726.03 |
| Weibull | 31.81 | 779.24 |
| Rayleigh | 24.53 | 600.97 |

**Interpretation**

- **Weibull:** Heavier right tail, so it predicts the highest mean power $\overline{v^3}$. Rare high-speed events lift the average.
- **GERD:** Smooth and stable. Slightly lower power than Weibull, but it fits the middle of the distribution better, so routine forecasts are often more dependable.
- **Rayleigh:** Simplest and thin-tailed. It underestimates power because it under-represents gusts and outliers.

As Graph 3 shows, Weibull offers the rosiest outlook for wind farms, and GERD provides a grounded middle ground.

In practice, these findings help policymakers and engineers choose the right model for resource assessment, turbine siting, and long-term planning. Bringing in hourly data and stronger hybrid models should sharpen the projections.

# DISCUSSION

Looking at Tokyo's wind data, the three models tell different stories. Weibull, because of its flexible form and heavy right tail, is good at picking up the occasional strong wind events that push the mean of $\boldsymbol{v^3}$. That's why it tends to predict the most power, and it explains why Weibull shows up so often in wind resource studies (Safari & Gasore, 2010; Serban et al., 2020; Parajuli, 2016).

GERD works differently. It mixes a Rayleigh base with a Gaussian term, which lets it fit the middle and moderate tails of the distribution. That matters in a place like Tokyo where terrain and buildings disrupt wind flow (Kondo et al., 2008; Yamashita, 1990). GERD usually comes in below Weibull on power estimates, but the forecasts are steadier and often feel more realistic day to day.

Rayleigh is the simplest. It runs fast but its thin tails mean it misses strong winds, so the power it predicts is low (Hennessey, 1978; Chattamvelli & Shanmugam, 2021). Monte Carlo simulations make the differences clear by generating synthetic series and showing how $\overline{\boldsymbol{v^3}}$ behaves. Here we worked with 21 years of monthly averages, but moving to hourly or daily data and bringing in terrain, elevation, and building density would sharpen things further (Kubota et al., 2008). Bottom line: Weibull gives the high-energy case, GERD gives a balanced urban fit, and Rayleigh is fine for quick scoping when detail isn't the goal.

**Applications**

GERD handles the center of a distribution and the odd spikes at the same time. That's why it works in real projects, not just on paper. In wind work it's useful on hilly or built-up sites for quick site checks, turbine layout, and local energy profiles. For grid ops it turns into probability forecasts that steady supply, guide demand response, and cut imbalance risk. In cities it maps street- and building-scale flow so designers can improve ventilation, avoid wind trouble spots, and place green zones smartly. The same idea carries to climate risk, pollution plumes, aerosol days, temperature swings, and even wave heights. With live data GERD can update its own parameters, help plan storage and dispatch, and plug into IoT or SCADA control. It's also a clear teaching example of a hybrid distribution, and it pairs well with LSTMs or transformers for time series. With GIS you can vary the parameters across space and build energy or risk maps for planning.

**Future Scope of GERD**

GERD can be extended in several ways. Turning the Gaussian part into a mixture lets the model handle multiple peaks, which is useful for sites with multimodal wind regimes. Embedding GERD into time-series tools like ARIMA, LSTMs, or transformers would make it work for real-time forecasting. Its scope isn't limited to wind either it can describe rainfall intensity, temperature swings, solar irradiance, or even ocean wave heights, wherever you have a central pattern with occasional anomalies. Linking parameters with GIS data allows the model to adjust to local microclimates, while online learning keeps it current as new observations arrive, supporting smart-grid operations and adaptive wind-farm control. Coupling GERD with physical atmosphere models could create hybrid forecasts, and stress-testing it under climate-change scenarios would show whether it holds up for long-term planning.

## CONCLUSION

We tested Weibull, Rayleigh, and a new Gaussian-Enhanced Rayleigh (GERD) on two decades of Tokyo wind data. Each model points to a different picture of power. Weibull catches the rare, fast winds and pushes energy estimates up. GERD balances the middle and the tails, so its forecasts come out steadier and often more believable in a dense city. Rayleigh runs quickly but misses the extremes and tends to understate power.

Monte Carlo runs showed how those statistical choices flow into engineering terms: the fit you pick changes the turbine output you expect. That makes the model choice practical, not just cosmetic. With hourly or daily data, and by building in seasonality and site details, forecasts can get sharper. There is also room for hybrids and adaptive tools ensembles, learning models, or online updating to track changing conditions.

## REFERENCES

1. Alodat, M. T., & Anagreh, Y. N. (2011). Durations distribution of Rayleigh process with application to wind turbines. Journal of Wind Engineering and Industrial Aerodynamics, 99(5), 651–657. https://doi.org/10.1016/j.jweia.2011.02.006
2. Bidaoui, H., El Abbassi, I., El Bouardi, A., & Darcherif, A. (2019). Wind speed data analysis using Weibull and Rayleigh distribution functions, case study: five cities northern Morocco. Procedia Manufacturing, 32, 786–793. https://doi.org/10.1016/j.promfg.2019.02.277
3. Chattamvelli, R., & Shanmugam, R. (2021). Rayleigh distribution. In Continuous Distributions in Engineering and the Applied Sciences–Part II (pp. 245–252). Springer. https://doi.org/10.1007/978-3-030-73589-0_9
4. Dokur, E., Ceyhan, S., & Kurban, M. (2022). Analysis of wind speed data using Finsler, Weibull, and Rayleigh distribution functions. Electrica, 22(1), 52–60. https://doi.org/10.5152/electrica.2022.21096
5. Gorla, R. S. R., Pallikonda, M. K., & Walunj, G. (2020). Use of Rayleigh distribution method for assessment of wind energy output in Cleveland–Ohio. Renewable Energy Research and Applications, 1(1), 11–18.
6. Hennessey, J. P., Jr. (1978). A comparison of the Weibull and Rayleigh distributions for estimating wind power potential. Wind Engineering, 2(3), 156–164. https://www.jstor.org/stable/43749453
7. Kondo, H., Tokairin, T., & Kikegawa, Y. (2008). Calculation of wind in a Tokyo urban area with a mesoscale model including a multi-layer urban canopy model. Journal of Wind Engineering and Industrial Aerodynamics, 96(10–11), 1655–1666. https://doi.org/10.1016/j.jweia.2008.02.050
8. Kubota, T., Miura, M., Tominaga, Y., & Mochida, A. (2008). Wind tunnel tests on the relationship between building density and pedestrian-level wind velocity: Development of guidelines for realizing acceptable wind environment in residential neighborhoods. Building and Environment, 43(10), 1699–1708. https://doi.org/10.1016/j.buildenv.2007.10.015
9. Ouahabi, M. H., Benabdelouahab, F., & Khamlichi, A. (2017). Analyzing wind speed data and wind power density of Tetouan city in Morocco by adjustment to Weibull and Rayleigh distribution functions. Wind Engineering, 41(3), 174–184. https://doi.org/10.1177/0309524X17707733
10. Parajuli, A. (2016). A statistical analysis of wind speed and power density based on Weibull and Rayleigh models of Jumla, Nepal. Energy and Power Engineering, 8(7), 271–282. https://doi.org/10.4236/epe.2016.87026
11. Pishgar-Komleh, S. H., Keyhani, A., & Sefeedpari, P. (2015). Wind speed and power density analysis based on Weibull and Rayleigh distributions (a case study: Firouzkooh county of Iran). Renewable and Sustainable Energy Reviews, 42, 313–322. https://doi.org/10.1016/j.rser.2014.10.015

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

12. Safari, B., & Gasore, J. (2010). A statistical investigation of wind characteristics and wind energy potential based on the Weibull and Rayleigh models in Rwanda. Renewable Energy, 35(12), 2874–2880. https://doi.org/10.1016/j.renene.2010.04.017
13. Serban, A., Paraschiv, L. S., & Paraschiv, S. (2020). Assessment of wind energy potential based on Weibull and Rayleigh distribution models. Energy Reports, 6, 250–267. https://doi.org/10.1016/j.egyr.2020.11.259
14. Yamashita, S. (1990). The urban climate of Tokyo. Geographical Review of Japan, Series B, 63(1), 98–107. https://doi.org/10.4157/grj1984b.63.98
15. Zambak, M. F., Cahyadi, C. I., Helmi, J., Sofie, T. M., & Suwarno, S. (2023). Evaluation and analysis of wind speed with the Weibull and Rayleigh distribution models for energy potential using three models. International Journal of Energy Economics and Policy, 13(2), 427–432. https://doi.org/10.32479/ijeep.13941