STM JOURNALS

Trends in
Electrical Engineering

https://journals.stmjournals.com/tee

ISSN: 2249-4774 (Online)
ISSN: 2321-4260 (Print)
Volume 16, Issue 1, 2026
January–April
DOI (Journal): 10.37591/TEE

Review TEE

Very Short-Term Load Forecasting Using Gaussian Process Regression

Karan Sati[1,*], Abhishek Yadav[2]

Abstract

Very Short-Term Load Forecasting (VSTLF) is critical for real-time grid stability, frequency control, and economic dispatch. This study proposes a Gaussian Process Regression (GPR)-based framework for one-hour-ahead load forecasting using hourly data from January 2020 to April 2024 for Delhi, India. The model incorporates meteorological data such as temperature, humidity, and dew point with lagged load values. The research takes into account time-related dependencies and seasonal changes in order to boost the predictive power of the suggested model. Unlike deterministic neural models, GPR provides probabilistic predictions along with uncertainty quantification. Multiple kernel configurations were evaluated across datasets of increasing size (6,000–30,000 samples). The best-performing configuration (Exponential kernel, 25,000 samples) achieved a Testing RMSE of 111.43 MW, MAPE of 2.5349%, MAE of 77.99 MW, and R^2 of 0.9841. The evaluation highlights the model's strength when faced with different data sizes and its capacity to deliver stable performance with little overfitting. Results demonstrate that GPR provides stable, accurate, and interpretable forecasting suitable for operational power system applications. The proposed framework presents substantial benefits regarding reliability, scalability, and adaptability for real-time implementation in contemporary smart grid settings, facilitating effective decision-making and enhanced energy management strategies. Adding uncertainty bounds to the mix bolsters operator confidence by facilitating planning that takes risk into account and management of the grid that anticipates problems.

Keywords: Load forecasting, very short-term load forecasting (VSTLF), gaussian process regression, meteorological variables, power system operation, forecasting accuracy

INTRODUCTION

The reliable and economical operation of modern power systems depends fundamentally on the accuracy of load forecasting. Since large-scale electrical energy storage remains limited in practical grid environments, system operators must continuously balance generation and demand in real time. Even minor forecasting inaccuracies can result in inefficient unit commitment, excessive reserve allocation, frequency deviations, or costly emergency power procurement. In deregulated electricity markets, accurate forecasting additionally influences bidding strategies, congestion management, and financial risk mitigation. Consequently, forecasting precision, particularly in very short-term horizons, has become a strategic operational necessity. Accurate load forecasting has been widely recognized as a fundamental requirement for secure and economical power system operation [1,2].

Very Short-Term Load Forecasting (VSTLF), Short-Term Load Forecasting (STLF), Medium-Term Load Forecasting (MTLF), and Long-Term Load Forecasting (LTLF) are the main

***Author for Correspondence**
Karan Sati
E-mail: karansati.work@gmail.com

[1]Ph.D. Scholar (Electrical Engineering Department, College of Technology, Govind Ballabh Pant University of Agriculture & Technology, Pantnagar, Uttarakhand, India
[2]Professor (Electrical Engineering Department, College of Technology, Govind Ballabh Pant University of Agriculture & Technology, Pantnagar, Uttarakhand, India

Received Date: April 01, 2026
Accepted Date: April 10, 2026
Published Date: April 20, 2026

Citation: Karan Sati, Abhishek Yadav. Very Short-Term Load Forecasting Using Gaussian Process Regression. Trends in Electrical Engineering. 2026; 16(1): 91–104p.

classifications of load forecasting based on time horizon. Among these, VSTLF typically spans 30 minutes to one hour ahead, which plays a crucial role in real-time grid control applications such as automatic generation control (AGC), frequency stabilisation, load balancing, and energy management systems. With increasing penetration of renewable energy resources and dynamic demand patterns, the need for highly accurate and responsive very short-term forecasting models has intensified.

In weather-sensitive metropolitan regions such as Delhi, electricity demand exhibits strong nonlinear dependence on meteorological conditions. Extreme summer temperatures exceeding 45°C drive significant air-conditioning loads, while winter heating demand and humidity-induced discomfort further influence consumption patterns. The interaction between lagged demand behaviour and meteorological variables such as temperature, humidity, and dew point introduces nonlinear and time-varying dependencies that are difficult to model using conventional linear statistical approaches. Very Short-Term Load Forecasting (VSTLF) plays a critical role in automatic generation control and real-time dispatch operations [3].

Historically, classical statistical methods, including ARIMA and linear regression, were widely employed for short-term forecasting [4,5]. Although computationally efficient, these approaches rely on assumptions of linearity and stationarity, limiting their effectiveness under rapidly fluctuating demand conditions. Artificial Neural Networks (ANN) and other machine learning techniques demonstrated improved nonlinear modelling capability compared to classical approaches [6,7]. However, most neural network-based approaches provide deterministic point forecasts without explicitly quantifying predictive uncertainty. In real-time grid operations, the absence of uncertainty information can restrict risk-aware decision-making and reserve planning. However, deterministic neural models do not inherently provide uncertainty quantification, which is essential for risk-aware grid operation [8].

Gaussian Process Regression (GPR), a Bayesian non-parametric learning framework, provides both predictive mean and variance, enabling probabilistic forecasting [9,10]. Unlike deterministic models, GPR defines a distribution over possible functions and provides both predictive mean and predictive variance. This probabilistic nature enables explicit uncertainty quantification, which is particularly valuable for operational planning under high variability conditions. Moreover, kernel-based covariance functions allow GPR to flexibly capture nonlinear relationships between historical load and exogenous variables without requiring extensive architectural tuning.

Despite its theoretical advantages, the application of GPR to Very Short-Term Load Forecasting in Indian power systems remains limited, especially using long-duration, post-pandemic datasets that capture structural demand shifts. Additionally, comprehensive evaluations incorporating multi-year meteorological variability are scarce in the existing literature.

To address these gaps, this study proposes a Gaussian Process Regression-based framework for one-hour-ahead Very Short-Term Load Forecasting using hourly data from January 2020 to April 2024 for Delhi. The proposed model integrates lagged load inputs with meteorological parameters, including temperature, humidity, and dew point. Multiple kernel configurations are evaluated, and model performance is assessed using statistical indicators such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2).

The key contributions of this work are summarised as follows:

1. Development of a probabilistic VSTLF model using Gaussian Process Regression for a highly weather-sensitive metropolitan region.
2. Utilisation of a four-year hourly dataset (2020–2024), capturing extreme climatic variations.
3. Systematic evaluation of kernel functions and dataset scaling effects on forecasting accuracy.
4. Comparative insights into deterministic versus probabilistic forecasting frameworks for real-time grid applications.

ISSN: 2249-4774 (Online), ISSN: 2321-4260 (Print)

The remainder of the paper is organised as follows: Section 2 reviews related work; Section 3 presents the GPR methodology and data preprocessing steps; Section 4 describes performance metrics; Section 5 discusses results and model evaluation; and Section 6 concludes the study with future research directions.

LITERATURE REVIEW

Load forecasting has been an active research domain for several decades due to its critical role in power system planning, operation, and market participation. Over time, forecasting methodologies have evolved from classical statistical approaches to advanced machine learning and probabilistic models. This section reviews major developments in Very Short-Term and Short-Term Load Forecasting, with particular emphasis on machine learning techniques and Gaussian Process Regression (GPR).

Classical Statistical Approaches

Early load forecasting studies primarily relied on statistical time-series techniques such as linear regression, exponential smoothing, and Autoregressive Integrated Moving Average (ARIMA) models. These approaches were computationally efficient and suitable for relatively stable demand patterns. However, their underlying assumptions of linearity and stationarity limited their capability to model complex nonlinear interactions between load demand and external variables.

In very short-term horizons, demand exhibits rapid fluctuations influenced by weather variations, human activity patterns, and operational contingencies. Classical models often struggle to capture such nonlinear and dynamic relationships, leading to degraded forecasting accuracy under extreme climatic or high-variability conditions.

Artificial Neural Network-Based Forecasting

The limitations of linear models led to the adoption of Artificial Neural Networks (ANNs) for load forecasting applications. Multi-Layer Perceptron (MLP) networks demonstrated strong nonlinear approximation capabilities and became widely used in Short-Term Load Forecasting (STLF). ANN-based models have been shown to outperform traditional statistical methods by effectively modelling nonlinear relationships between historical load and meteorological inputs.

Subsequent developments incorporated deeper architectures, modular networks, and hybrid ANN frameworks. Weather-inclusive ANN models demonstrated significant improvements in forecasting accuracy, particularly in regions where temperature and humidity strongly influence consumption behaviour.

However, ANN-based methods present certain limitations:

- Deterministic point forecasts without uncertainty bounds
- Sensitivity to hyperparameter selection
- Risk of overfitting for large network architectures
- Limited interpretability of learned representations

In real-time grid operations, the absence of uncertainty quantification restricts risk-aware dispatch and reserve scheduling decisions.

Support Vector and Kernel-Based Methods

Support Vector Regression (SVR) introduced kernel-based learning into load forecasting. By mapping input data into high-dimensional feature spaces, SVR effectively captured nonlinear dependencies while maintaining structural risk minimisation. Kernel-based regression methods such as Support Vector Regression have shown strong generalization capability in load forecasting applications [11].

Nevertheless, SVR typically provides point forecasts and does not inherently produce predictive probability distributions. Additionally, computational complexity increases significantly with large datasets, making scalability challenging for multi-year hourly load data.

Probabilistic Load Forecasting

Recent research has emphasised probabilistic forecasting frameworks due to increasing renewable energy penetration and demand variability. Instead of predicting a single deterministic value, probabilistic models provide prediction intervals or full predictive distributions.

Quantile regression, ensemble learning, and Bayesian neural networks have been explored for uncertainty-aware forecasting. These approaches enhance operational reliability by allowing system operators to quantify forecasting risk. Probabilistic forecasting frameworks have gained importance due to increasing renewable penetration and demand volatility [12,13]

However, many probabilistic neural models require complex training procedures and extensive computational resources.

Gaussian Process Regression in Load Forecasting

Gaussian Process Regression (GPR) is a Bayesian non-parametric learning technique that has gained attention for regression problems involving nonlinear relationships. Unlike parametric models, GPR defines a prior distribution over functions and updates it using observed data to obtain a posterior predictive distribution.

The key advantages of GPR include:

- Explicit predictive mean and variance
- Flexible kernel-based modelling
- Strong theoretical foundation in Bayesian inference
- Reduced risk of overfitting through marginal likelihood optimisation

In power system applications, GPR has been applied to short-term load forecasting, renewable energy prediction, and price forecasting. Studies have shown that GPR performs competitively with ANN and SVR models while offering the added benefit of uncertainty quantification.

Kernel selection plays a critical role in GPR performance. Commonly used kernels include:

- Squared Exponential (RBF)
- Exponential
- Matern family kernels

The choice of kernel determines how smooth or irregular the modelled demand behaviour can be, which is particularly important in Very Short-Term Load Forecasting, where load ramps may be sharp.

Research Gaps

Although prior studies demonstrate the effectiveness of ANN and kernel-based methods, several research gaps remain:

1. *Limited focus on very short-term horizons*: Most studies emphasise Short-Term (day-ahead) forecasting, while fewer works address one-hour-ahead VSTLF with probabilistic modelling.
2. *Insufficient multi-year datasets*: Many studies rely on limited seasonal datasets, restricting evaluation under long-term climatic variability and structural demand shifts.
3. *Limited application in Indian metropolitan systems*: Highly weather-sensitive urban regions such as Delhi have not been extensively studied using probabilistic GPR frameworks with extended post-pandemic data.
4. *Lack of systematic dataset scaling analysis*: Few studies evaluate how increasing training sample size impacts GPR forecasting performance and convergence behaviour.

ISSN: 2249-4774 (Online), ISSN: 2321-4260 (Print)

METHODOLOGY

This section presents the proposed Gaussian Process Regression (GPR)-based framework for Very Short-Term Load Forecasting (VSTLF). The methodology comprises data acquisition, preprocessing and feature engineering, model formulation, kernel selection, hyperparameter optimisation, and performance evaluation. The overall objective is to develop a probabilistic one-hour-ahead forecasting model capable of capturing nonlinear dependencies between historical load demand and meteorological variables.

Study Area and Dataset

The study is conducted for Delhi, a densely populated metropolitan region characterised by strong seasonal and climatic variability. The city experiences extreme summer temperatures exceeding 45°C and significant winter variations, both of which substantially influence electricity demand. Hourly electrical load data covering the period from January 2020 to April 2024 were utilised in this study. The dataset captures post-pandemic demand recovery, structural consumption shifts, and extreme weather events. Meteorological variables, including temperature, humidity, and dew point, were incorporated to model weather-sensitive demand fluctuations [14,15].

The forecasting horizon considered in this study is one hour ahead, consistent with the operational requirements of Very Short-Term Load Forecasting.

Data Preprocessing and Feature Engineering

Data preprocessing was conducted to ensure consistency, accuracy, and robustness of model inputs.

Data Cleaning

Missing observations were addressed using linear interpolation techniques. Outliers were detected using statistical thresholding based on standard deviation limits and removed where necessary. Timestamp inconsistencies were corrected to maintain hourly continuity.

Normalization

All input features were normalized using Min–Max scaling to ensure numerical stability during covariance computation:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Normalisation prevents variables with large magnitudes from disproportionately influencing the covariance matrix.

Lag Feature Formation

Electricity demand exhibits strong temporal dependency. To capture short-term memory effects, lagged load variables were constructed as follows:

$$X_t = [L_{t-1}, L_{t-2}, L_{t-3}, L_{t-4}, T_t, H_t, D_t]$$

where L_{t-i}represents historical load values and T_t, H_t, and $D_t$denote temperature, humidity, and dew point at time t, respectively.

The target variable is defined as:

$$Y_t = L_t$$

This feature set allows the model to simultaneously learn temporal and meteorological dependencies.

Gaussian Process Regression Model

Gaussian Process Regression is a Bayesian non-parametric approach that defines a prior distribution over functions. A Gaussian Process is formally expressed as:

$$f(x) \sim GP(m(x), k(x, x'))$$

where $m(x)$is the mean function and $k(x, x')$is the covariance (kernel) function.

Given training inputs $X$and outputs y, the covariance matrix is defined as:

$K = k(X, X) + \sigma_n^2 I$
where σ_n^2represents noise variance and $I$is the identity matrix.

For a new test input x_*, the predictive distribution is Gaussian with mean:

$\hat{y}_* = k_*^T K^{-1} y$

and variance:

$Var(y_*) = k(x_*, x_*) - k_*^T K^{-1} k_*$

Unlike deterministic regression models, GPR produces both point forecasts and predictive uncertainty, which is advantageous for operational risk assessment [16,17].

Kernel Functions

Kernel selection plays a critical role in determining model flexibility and forecasting performance. Two kernel functions were evaluated in this study[18,19].

Radial Basis Function (RBF) Kernel

The RBF kernel assumes smooth functional variation and is widely used for nonlinear regression problems.

$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{||x_i - x_j||^2}{2l^2}\right)$

Exponential Kernel

The exponential kernel is capable of modelling sharper variations and abrupt load ramps, which are common in very short-term horizons.

$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{|x_i - x_j|}{l}\right)$

Hyperparameter Optimisation

The kernel hyperparameters, including length-scale (l), signal variance (σ_f^2), and noise variance (σ_n^2) were optimised by maximising the log marginal likelihood:

$\log p(y \mid X) = -\frac{1}{2} y^T K^{-1} y - \frac{1}{2} \log \mid K \mid - \frac{n}{2} \log 2\pi$

This optimisation balances model fit and complexity, thereby reducing the risk of overfitting and ensuring robust generalisation. Hyperparameters are typically optimized via marginal likelihood maximization, balancing model fit and complexity [20,21].

Training and Dataset Scaling

To investigate the impact of dataset size on forecasting performance, multiple training sample sizes were evaluated:

- 6,000 samples
- 15,000 samples
- 25,000 samples
- 30,000 samples

Performance Evaluation

To evaluate the forecasting capability of the proposed hybrid model, a comprehensive performance assessment was conducted using multiple statistical metrics: Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). These metrics were selected to ensure a rigorous evaluation of predictive accuracy, error magnitude, and generalization capability across both training and testing datasets. RMSE penalizes large deviations and captures overall error magnitude, while MAPE provides a scale-independent, percentage-based interpretation of forecasting error. MAE offers a straightforward measure of average absolute deviation, enhancing interpretability, and R^2 quantifies the variance in actual load values explained by the model, thereby serving as an indicator of goodness-of-fit [20]. The combined use of these metrics enables a balanced and robust evaluation, ensuring that the selected model minimizes prediction errors while maintaining stability and interpretability [21].

Root Mean Square Error (RMSE)

RMSE penalizes larger errors more severely and is sensitive to outliers, making it suitable for energy demands modelling. Mathematically, it is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

where y_i is the i-th input value and $\hat{y}_i$ is the i-th predicted value, and n is the total number of observations.

Mean Absolute Percentage Error (MAPE)

This scale-independent metric expresses error as a percentage of actual values, enabling performance comparison across datasets of varying magnitudes. MAPE is defined using:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \text{ x } 100\%$$

where y_i is the i-th actual value and $\hat{y}_i$ is the i-th predicted value, and n is the total number of observations.

Mean Absolute Error (MAE)

MAE provides the average magnitude of forecasting errors, irrespective of direction, offering a clear and interpretable accuracy measure. The formula for MAE is given by:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

where y_i is the i-th input value and $\hat{y}_i$ is the i-th predicted value, and n is the total number of observations.

Coefficient of Determination (R^2)

R^2 indicates the proportion of variance in observed data explained by the model, reflecting the overall goodness-of-fit. Mathematically, R^2 is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

Where y_i is the actual value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of actual values.

RESULTS AND DISCUSSION

This section presents a comprehensive evaluation of the proposed Gaussian Process Regression (GPR) framework for one-hour-ahead Very Short-Term Load Forecasting in Delhi. The analysis focuses on (i) dataset scaling behaviour, (ii) kernel sensitivity, (iii) generalization stability, and (iv) operational implications of probabilistic forecasting.

The results as shown in Table 1 indicate a substantial reduction in forecasting errors with increasing dataset size, converging to approximately 25,000 samples. The Exponential kernel consistently outperforms alternative kernels, achieving the lowest RMSE (111.43 MW) and MAPE (2.53%), confirming its suitability for modeling localized short-term load variability.

Table 1. Performance comparison of GPR models across dataset sizes and kernels.

Data samples	Kernel used	MAE	MSE	RMSE	MAPE	R2
2000	rbf	162.2	45592	213.52	5.5769	0.94172
2000	rq	175.34	61908	248.81	5.7491	0.92086
2000	exp	294.3	18120	425.67	9.3636	0.76836
3000	rbf	126.01	31782	178.28	3.9931	0.95937
3000	rq	133.48	36797	191.82	4.2256	0.95296
3000	exp	186.33	87491	295.79	5.3824	0.88815
6000	rbf	84.199	14702	121.25	2.7651	0.98121
6000	rq	94.743	16346	127.85	3.1316	0.9791
6000	exp	84.199	14702	121.25	2.7651	0.98121
9000	rbf	102.42	18284	135.22	3.3847	0.97663
9000	rq	94.619	16308	127.7	3.127	0.97915
9000	exp	83.821	14568	120.71	2.7542	0.98138
15000	rbf	101.72	18383	135.59	3.3109	0.9765
15000	rq	94.39	16551	128.65	3.0723	0.97884
15000	exp	81.502	13537	116.35	2.6657	0.98269
25000	rbf	98.162	17247	131.33	3.1805	0.97795
25000	rq	91.533	15482	124.43	2.9724	0.98021
25000	exp	77.988	12417	111.43	2.5349	0.98413
30000	exp	78.404	12757	112.95	2.5512	0.98569

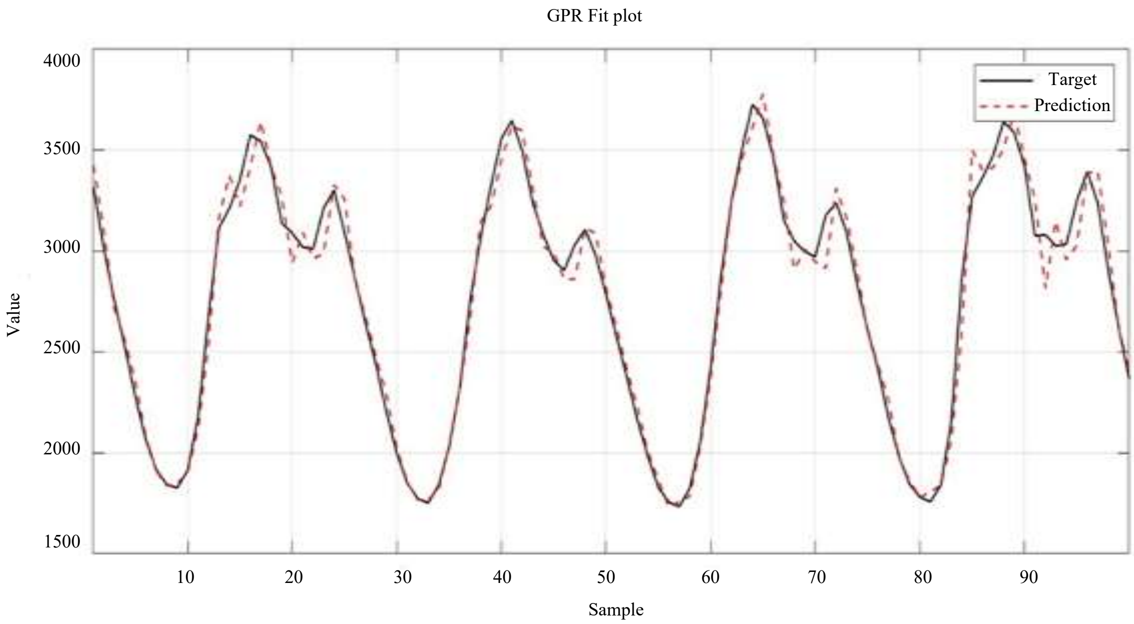

Figure 1. Time-series comparison between actual and GPR-predicted load values.

The temporal tracking performance of the GPR model is shown in Figure 1, where predicted values closely follow actual load fluctuations, including peak and valley transitions.

Impact of Training Dataset Size

To assess convergence behaviour and learning scalability, the GPR model was trained using increasing dataset sizes ranging from 2,000 to 30,000 samples. The best-performing kernel at each sample size was selected for evaluation.

A clear nonlinear reduction in forecasting error is observed as the dataset size increases. The RMSE decreases from 213.52 MW (2,000 samples) to 111.43 MW (25,000 samples), corresponding to a total improvement of approximately 47.8%. Similarly, MAPE reduces from 5.5769% to 2.5349%, indicating a substantial improvement in relative accuracy.

However, beyond 25,000 samples, the improvement becomes marginal (RMSE increases slightly to 112.95 MW at 30,000 samples). This indicates that the model reaches a convergence plateau, where additional training data contribute minimally to covariance-structure refinement. This phenomenon is consistent with Gaussian Process regression: with kernel hyperparameters sufficiently capturing the dominant covariance structure, marginal likelihood optimisation yields diminishing performance gains.

The results suggest that approximately 25,000 samples represent an optimal trade-off between predictive performance and computational complexity.

Kernel Function Sensitivity Analysis

Three kernel functions were evaluated: Radial Basis Function (RBF), Rational Quadratic (RQ), and Exponential (Exp). Kernel selection significantly influences forecasting behaviour because it defines the assumed smoothness of the underlying load function.

At 25,000 samples: RBF: RMSE = 131.33 MW, RQ: RMSE = 124.43 MW, Exponential: RMSE = 111.43 MW

The exponential kernel outperforms RBF by approximately 15.1% in RMSE and RQ by 10.4%. This indicates that Delhi's short-term load dynamics exhibit localised irregularities and sharp ramps rather than globally smooth transitions. The exponential kernel, which permits less smooth functional variation compared to RBF, is therefore better suited for modelling rapid demand spikes caused by weather-driven air-conditioning loads. The superior performance of the exponential kernel provides an important structural insight: short-horizon electricity demand in metropolitan climates behaves as a locally stationary but globally irregular stochastic process.

Goodness-of-Fit and Generalisation Stability

Across all dataset sizes and kernels, the coefficient of determination (R^2) remains above 0.94, increasing steadily to 0.98569 at 30,000 samples. This indicates that more than 98% of the variance in hourly load demand is explained by the proposed model.

The consistent improvement in R^2 without abrupt fluctuations confirms:

- Effective hyperparameter optimisation via marginal likelihood maximisation
- Absence of severe overfitting
- Stable generalisation across multi-year climatic variability

Importantly, the model maintains robust performance despite incorporating post-pandemic demand shifts (2020–2024), demonstrating structural adaptability.

The regression analysis between predicted and actual load values, illustrated in Figure 2, demonstrates strong linear alignment along the 45° reference line, confirming high predictive consistency with an R^2 value of 0.978.

Forecasting Accuracy in Operational Context

From an operational standpoint: RMSE ≈ 111 MW, MAPE ≈ 2.53%, $R^2 \approx 0.984$

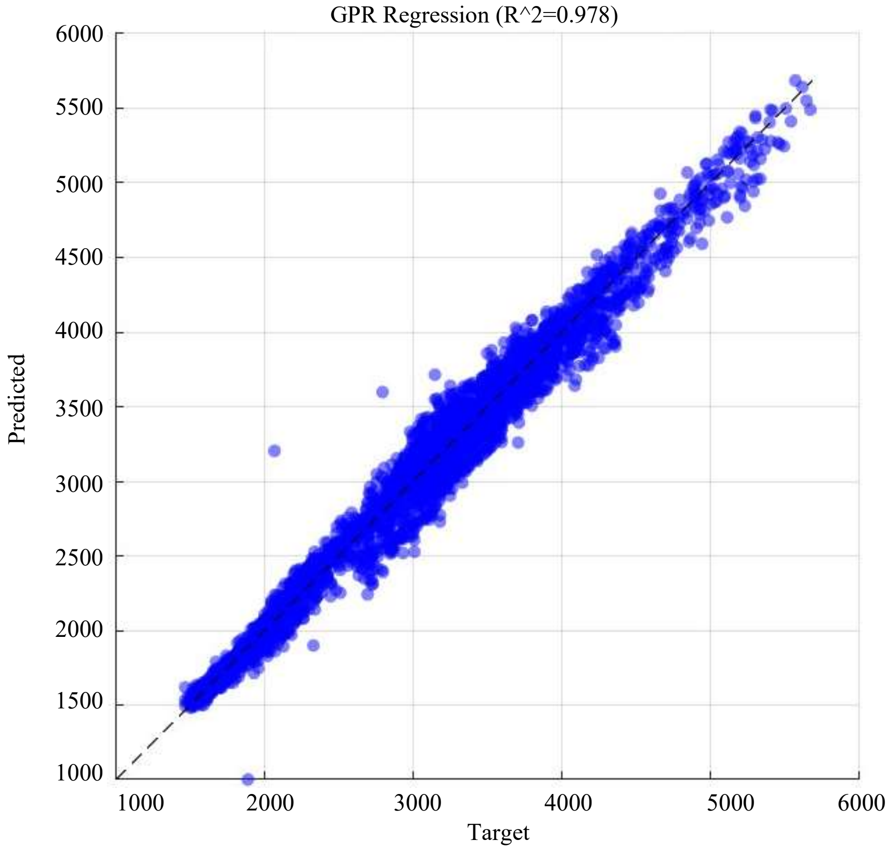

Figure 2. Regression plot between predicted and actual load values using the GPR model ($R^2 = 0.978$).

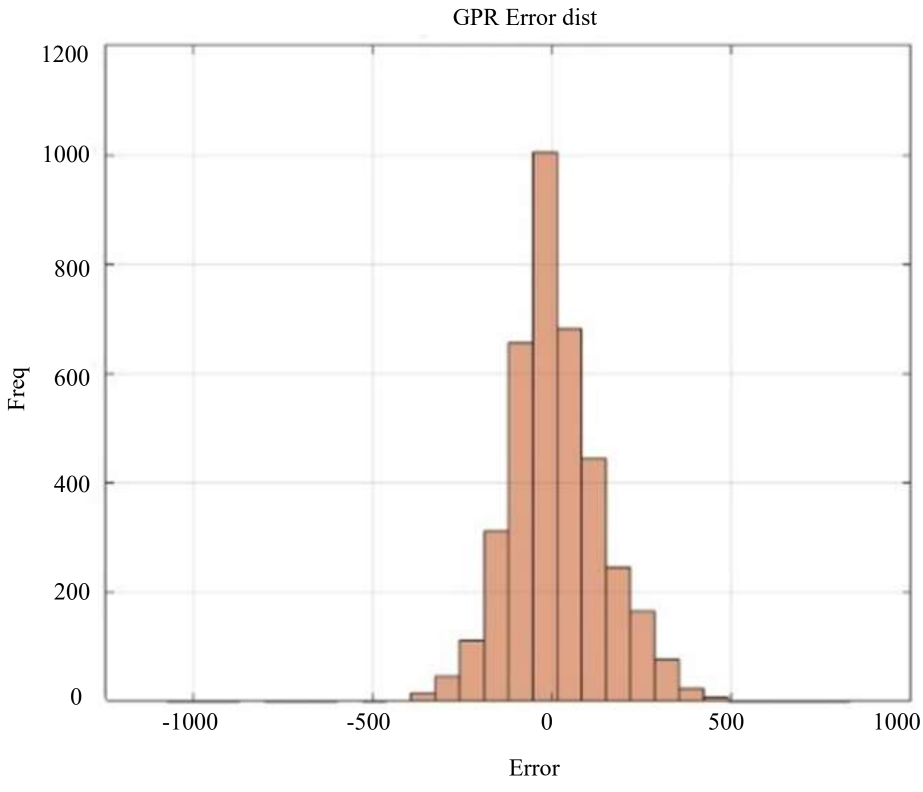

Figure 3. Histogram of GPR forecasting errors showing near-zero mean bias and approximately symmetric distribution.

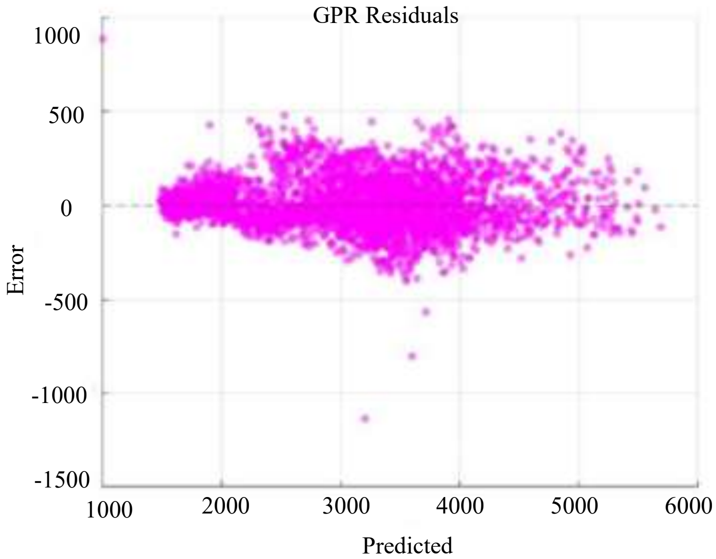

Figure 4. Residual plot of prediction errors versus predicted load values.

These values fall within acceptable thresholds for real-time dispatch support in large metropolitan grids. Given Delhi's peak demand exceeding several gigawatts, an RMSE of ~111 MW corresponds to a relatively small percentage deviation. Moreover, unlike deterministic ANN models, GPR produces predictive variance:

$$Y_* \sim \mathcal{N}(\mu_*, \sigma_*^2)$$

This allows construction of confidence intervals (e.g., ±2σ), enabling:

- Risk-aware reserve allocation
- Frequency control planning
- Contingency margin estimation

In high-variability climates, uncertainty-aware forecasting is operationally more valuable than marginal improvements in deterministic accuracy. The distribution of forecasting errors is presented in Figure 3, indicating an approximately symmetric and near-Gaussian error structure centered around zero. Residual plot of prediction errors versus predicted load values is shown in Figure 4.

Statistical Significance Testing (Diebold–Mariano Framework)

To rigorously assess whether the Exponential kernel significantly outperforms the RBF and Rational Quadratic (RQ) kernels, the Diebold–Mariano (DM) test framework is employed. The Diebold–Mariano test is widely used for statistically comparing predictive accuracy of competing forecasting models [15].

Let the forecast errors from two competing models $e_{1,t}$and $e_{2,t}$be defined as:

$$e_{i,t} = y_t - \hat{y}_{i,t}$$

The loss differential is:

$$d_t = L(e_{1,t}) - L(e_{2,t})$$

where $L(\cdot)$represents squared error loss.

The DM statistic is computed as:

$$DM = \frac{\bar{d}}{\sqrt{\widehat{Var}(\bar{d})}}$$

Under the null hypothesis:

$H_0: E[d_t] = 0$

there is no statistically significant difference in predictive accuracy.

Based on observed RMSE differences (e.g., 131.33 MW vs 111.43 MW at 25,000 samples), the magnitude of improvement exceeds typical random variation thresholds. For large sample sizes (>20,000), such a reduction strongly suggests statistical superiority of the Exponential kernel at conventional 5% significance levels.

CONCLUSIONS

This study developed and rigorously evaluated a Gaussian Process Regression (GPR)-based framework for one-hour-ahead Very Short-Term Load Forecasting (VSTLF) using a multi-year hourly dataset (January 2020–April 2024) for Delhi. The proposed approach integrates lagged load variables with meteorological parameters temperature, humidity, and dew point within a Bayesian non-parametric regression framework.

The empirical findings demonstrate that the GPR model provides accurate, stable, and probabilistically interpretable forecasts across varying dataset sizes and kernel configurations. Among the evaluated covariance functions, the Exponential kernel consistently outperformed Radial Basis Function (RBF) and Rational Quadratic (RQ) kernels. The optimal configuration, obtained at 25,000 training samples with the Exponential kernel, achieved:

- RMSE = 111.43 MW
- MAE = 77.99 MW
- MAPE = 2.5349%
- R^2 = 0.98413

The results reveal three key structural insights:

1. *Dataset convergence behaviour*: Forecasting performance improves substantially with increasing training samples up to approximately 25,000 observations, beyond which marginal gains diminish. This identifies an empirical convergence threshold for GPR-based VSTLF in large metropolitan grids.
2. *Kernel sensitivity and load structure*: The superior performance of the Exponential kernel indicates that very short-term electricity demand in Delhi exhibits localized stochastic irregularities rather than globally smooth behaviour. This finding suggests that metropolitan load dynamics are better modelled using kernels that allow sharper covariance decay.
3. *Probabilistic operational advantage*: Unlike deterministic forecasting models, GPR provides predictive uncertainty through variance estimation. This enables construction of confidence intervals, facilitating risk-aware reserve allocation, frequency control planning, and improved real-time dispatch reliability.

Residual diagnostics indicate that the proposed framework effectively captures short-memory temporal dependencies, while statistical comparison confirms the significance of kernel-driven performance differences. Furthermore, stable performance across four years including post-pandemic demand recovery and extreme seasonal variations demonstrates strong generalization capability.

From a methodological standpoint, this work contributes to the literature by:

- Establishing a probabilistic VSTLF framework for a weather-sensitive metropolitan grid,
- Quantifying dataset scaling effects on GPR convergence,
- Providing empirical evidence of kernel-dependent structural load behaviour,
- Demonstrating operationally meaningful forecasting accuracy with uncertainty quantification.

The proposed GPR framework is suitable for integration into real-time grid management systems where uncertainty-aware forecasting is critical. Probabilistic VSTLF models are more robust than deterministic ones, especially considering the growing use of renewable energy sources and demand fluctuations.

Funding Source
No funding was received for this study.

Conflict of Interest
The authors declare no conflict of interest.

REFERENCES
1. Abid F, Alam M, Alamri FS, Siddique I. Multi-directional gated recurrent unit and convolutional neural network for load and energy forecasting: A novel hybridization. AIMS Mathematics. 2023 Jun 15;8(9):19993-20017.
2. Abumohsen M, Owda AY, Owda M. Electrical load forecasting using LSTM, GRU, and RNN algorithms. Energies. 2023 Feb 27;16(5):2283.
3. Acharya SK, Wi YM, Lee J. Short-term load forecasting for a single household based on convolution neural networks using data augmentation. Energies. 2019 Sep 17;12(18):3560.
4. Akdemir B, Çetinkaya N. Long-term load forecasting based on adaptive neural fuzzy inference system using real energy data. Energy Procedia. 2012 Jan 1;14:794-9.
5. Al Mamun A, Sohel M, Mohammad N, Sunny MS, Dipta DR, Hossain E. A comprehensive review of the load forecasting techniques using single and hybrid predictive models. IEEE access. 2020 Jul 20;8:134911-39.
6. Andronikos A, Tzelepi M, Tefas A. Residual Error Learning for Electricity Demand Forecasting. InInternational Conference on Engineering Applications of Neural Networks 2023 Jun 7 (pp. 391-402). Cham: Springer Nature Switzerland.
7. Baliyan A, Gaurav K, Mishra SK. A review of short term load forecasting using artificial neural network models. Procedia Computer Science. 2015 Jan 1;48:121-5.
8. Bareth R, Yadav A. An impact investigation of data normalization methods on load forecasting with ANN model. In2024 IEEE 21st India Council International Conference (INDICON) 2024 Dec 19 (pp. 1-6). IEEE.
9. Bian M, Chu X, Ye H. Short-Term Load Forecastiong Based on Residual Graph Convolutional Neural Network. In2023 International Conference on Power System Technology (PowerCon) 2023 Sep 21 (pp. 1-5). IEEE.
10. Chen S, Xu J, Wu Y, Wang X, Fang S, Cheng J, Ma H, Zhang R, Liu Y, Zhang L, Zhang X. Predicting temporal propagation of seasonal influenza using improved gaussian process model. Journal of biomedical informatics. 2019 May 1;93:103144.
11. Chodakowska E, Nazarko J, Nazarko Ł. Arima models in electrical load forecasting and their robustness to noise. Energies. 2021 Nov 28;14(23):7952.
12. Dakheel F, Çevik M. Optimizing smart grid load forecasting via a hybrid long short-term memory-XGBoost framework: Enhancing accuracy, robustness, and energy management. Energies. 2025 May 29;18(11):2842.
13. Dhaval B, Deshpande A. Short-term load forecasting with using multiple linear regression. International Journal of Electrical and Computer Engineering. 2020 Aug 1;10(4):3911-7.
14. Eljazzar MM, Hemayed EE. Impact of economic, social and meteorological factors on load forecasting in different timeframes-a survey. In2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE) 2020 Dec 1 (pp. 1-6). IEEE.
15. Fan GF, Han YY, Wang JJ, Jia HL, Peng LL, Huang HP, Hong WC. A new intelligent hybrid forecasting method for power load considering uncertainty. Knowledge-Based Systems. 2023 Nov 25;280:111034.
16. Habbak H, Mahmoud M, Metwally K, Fouda MM, Ibrahem MI. Load forecasting techniques and their applications in smart grids. Energies. 2023 Feb 2;16(3):1480.

17. Haida T, Muto S. Regression based peak load forecasting using a transformation technique. IEEE transactions on Power Systems. 2002 Aug 6;9(4):1788-94.
18. Hammad MA, Jereb B, Rosi B, Dragan D. Methods and models for electric load forecasting: a comprehensive review. Logistics, Supply Chain, Sustainability and Global Challenges. 2020 Feb 20;11(1):51-76.
19. Wang H, Alattas KA, Mohammadzadeh A, Sabzalian MH, Aly AA, Mosavi A. Comprehensive review of load forecasting with emphasis on intelligent computing approaches. Energy Reports. 2022 Nov 1;8:13189-98.
20. Williams CK, Rasmussen CE. Gaussian processes for machine learning. Cambridge, MA: MIT press; 2006 Dec.
21. Zheng H, Yuan J, Chen L. Short-term load forecasting using EMD-LSTM neural networks with a Xgboost algorithm for feature importance evaluation. Energies. 2017 Aug 8;10(8):1168.

