Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

Manaus, v.12 n.57, p. 375-388. January/February, 2026.

DOI: https://doi.org/10.5935/jetia. v12i57.3015

ISSN ONLINE: 2447-0228

RESEARCH ARTICLE OPEN ACCESS

AN EVALUATION OF DIRECT AND INDIRECT STRATEGIES FOR HIERARCHICAL NET-LOAD FORECASTING

Oyebola Olasupo*[1], Temitope Raphael Ayodele[2], Ayodeji Samson Olatunji Ogunjuyigbe[3]

[1, 2, 3]Department of Electrical and Electronic Engineering, University of Ibadan, Nigeria.

[1]https://orcid.org/0000-0003-0753-5721, [2]https://orcid.org/ 0000-0001-7789-934X, [3]https://orcid.org/0000-0002-0387-7380

Email: *oolasupo8962@stu.ui.edu.ng, tayodele2001@yahoo.com, s.ogunjuyigbe@ui.edu.ng

ARTICLE INFO

Article History
Received: November 27, 2025
Revised: December 10, 2025
Accepted: January 1, 2026
Published: January 31, 2026

Keywords:
Netload,
Forecasting,
Reconciliation,
Solar,
Time Series.

ABSTRACT

Power systems require accurate and coherent net-load forecasts across all levels of the grid hierarchy. A key methodological choice is between a direct strategy (forecasting net-load holistically) and an indirect strategy (forecasting gross load and renewable generation separately). This paper provides the first systematic evaluation of these two approaches within a hierarchical net-load forecasting framework, including the subsequent impact of statistical forecast reconciliation. Using load data from a three-level electricity network and simulated solar PV generation, this study compares the performance of both strategies using neural network auto-regression forecasting model and five reconciliation techniques. The findings show that the direct forecasting strategy yields statistically significant and more accurate forecasts across the hierarchy. However, the direct approach higher accuracy is coupled with over-forecasting bias. The analysis of reconciliation techniques revealed that while optimal methods improved the direct strategy's accuracy, they failed to correct its bias. Conversely, the simple Bottom-Up method, when applied to the indirect strategy, provided the least biased, though less accurate, overall forecast. The study concludes that while the direct approach is more robust for maximizing forecast accuracy, the final choice of strategy and reconciliation method must be aligned with specific operational objectives.

I. INTRODUCTION

Accurate forecasting of future electricity demand is crucial for efficient generation, transmission and distribution [1-3]. Net-load is defined as the electricity demand minus renewable energy sources (RES) generation [4]. The dual uncertainties from demand and supply variability complicate net-load forecasting (NLF) [1], making it a challenge in the transition to decarbonised power grids [5-7]. To improve the accuracy of NLF, studies have explored various methodological approaches. These include direct NLF (forecasting net-load holistically) and indirect NLF (forecasting gross load and RES generation separately, then adding to form NLF) strategies [6], [8], application of traditional statistical models, Machine Learning (ML) and Deep Learning (DL) architectures [9], [10], and the use of pre-processing techniques like signal decomposition and feature engineering to handle the non-linearity inherent in net-load data [11], [12]. The integration of RES such as solar PV into power grids across the globe necessitates accurate net-load forecasts at various granularities, from individual feeders and substations to the total system level, for effective planning, scheduling and real-time operations.

The existing body of research has extensively addressed the challenge of improving point net-load forecast accuracy at a single level of aggregation. Direct NLF, which forecasts the net load as a single time series, is widely adopted due to its simplicity and applicability in scenarios with limited observability of behind-the-meter (BTM) RES. In contrast, indirect (or additive) forecasting involves separately predicting load and RES generation before computing the net load. Existing literature on direct versus indirect strategies in net load forecasting reports varied outcomes, often due to evaluations limited to a single aggregation level or specific contexts such as microgrids or high renewable penetration systems. For example, Alipour *et al.* [11] integrated deep neural networks with wavelet transforms and found the indirect strategy more accurate, with improvements in metrics like mean absolute percentage error when using transformed historical inputs.

[13] applied LSTM models in microgrids with solar and wind, reporting lower MAE, MSE, RMSE, and NRMSE for the indirect method compared to direct. [14] proposed a disaggregation based additive framework using dictionary learning and LSTM, achieving RMSE reductions of 9.41% to 16.01% and MAE reductions of 9.34% to 15.24% over direct forecasts in neighbourhoods with high BTM PV and low observability. [15] used unsupervised contrastive disaggregation followed by models like DeepAR and TFT, demonstrating superior accuracy for the additive approach over direct net load forecasting. [16] tested machine learning models with and without BTM PV data, noting slight advantages for indirect but minimal differences overall. [17] compared additive (separate load and solar) versus integrated (solar as input to net load model) approaches, finding the integrated method superior due to cointegration of errors. Conversely, some studies support direct strategies.

[18] employed BNNs in renewable microgrids and reported lower NRMSE (3.98%) for direct compared to indirect (4.17%). [19] compared aggregation levels, identifying partially aggregated (separate total demand and supply) as optimal over fully direct or disaggregated. These variations stem from differences in models, data granularity, renewable types, and metrics, with limited attention to hierarchical structures or bias alongside accuracy. While direct and indirect NLF debate is well-documented in the literature, are well-investigated, their application to hierarchical time series, which naturally represent multi-level power systems, remains unexplored in NLF literature. When forecasts are generated independently for each level of the grid, the resulting set of forecasts is almost certainly incoherent which means that the sum of the forecasts for lower-level nodes (e.g., feeders) will not be equal to the forecast for their parent node (e.g., a substation). This incoherence creates significant operational challenges.

The transmission system operator, relying on an aggregate forecast, may make dispatch decisions that conflict with the needs identified by a distribution operator, who is relying on substation-level forecasts. Incoherent forecasts also lead to suboptimal scheduling of ancillary services, energy storage and demand response, as the perceived need at the aggregate level does not align with the sum of needs at the local level. Forecast reconciliation techniques ensure this coherence by adjusting independent base forecasts [20]. The primary contribution of this paper is an evaluation of direct and indirect strategies for hierarchical net-load forecasting (HNLF) and reconciliation. Unlike prior studies which are often limited to a single model or aggregation level, this work compares these two approaches for a hierarchical electricity network with solar photo-voltaic (PV) penetration, using three forecasting models and five statistical reconciliation techniques. The findings of this study is important for identifying effective forecasting strategies in RES-dominated power systems.

III. MATERIALS AND METHODS

I. Data: The primary datasets used in this study are the gross electricity load data and exogenous data. The gross load data for this study was obtained from an electricity network which covers a significant part of a state in the south-western part of Nigeria. Nine distribution feeders at 33kV voltage level are energised via the secondary side of three 132/33kV transformers at a 132/33kV substation. The hierarchical structure of the network is as shown in Figure 1. Historical hourly electricity load data for the feeders are acquired between 00:00 hours of 1st January 2022 to 24:00 hours of 31st December 2024. The raw data is validated and cleaned, while missing values and outliers are addressed. After cleaning, let $\mathbf{y}_t = (y_{1,t}, y_{2,t}, \dots, y_{m,t})^\top \in \mathbb{R}^m, t = 1,2, \dots, T$ be the vector of hourly electricity load observations at the m bottom-level nodes (e.g., feeders) over a horizon of T hours. Then, the complete multi-level load vector is $\mathbf{Y}_t = \mathbf{S}\mathbf{y}_t, \mathbf{Y}_t \in \mathbb{R}^n$, where $n \geq m$ and $\mathbf{S} \in \{0,1\}^{n \times m}$ denote the summing matrix that maps the m bottom-level series to n upper level series (such as transformers and substations).

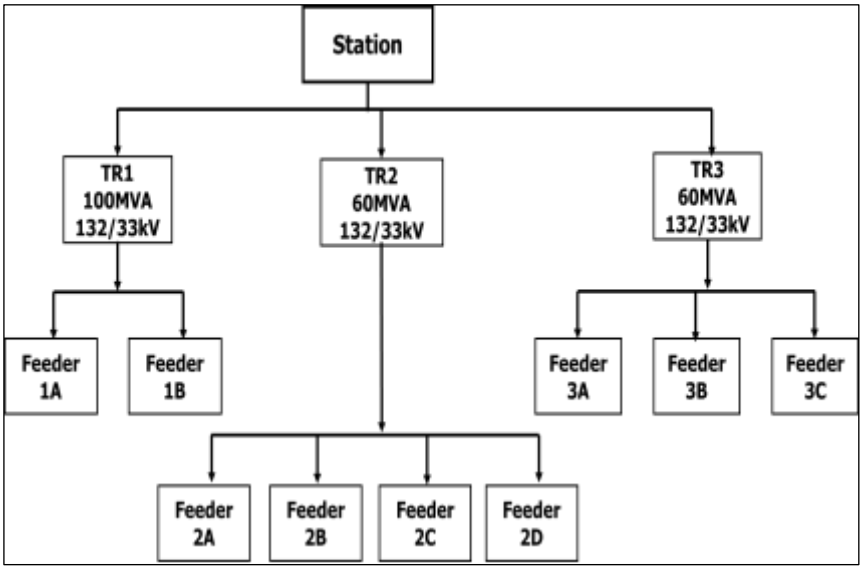

Figure1: Hierarchical Structure of the Electricity Network.
Source: Authors, (2026).

Exogenous data in this study encompasses external variables that influence net load calculations. The exogenous data include hourly meteorological and temporal datasets for the specific geographical location covering the entire historical period, $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \dots, x_{p,t})^\top \in \mathbb{R}^p$). The Global Horizontal Irradiance (GHI, measured in W/m²), was obtained from the NASA POWER database (https://power.larc.nasa.gov/) , providing hourly values that quantify the solar radiation incident on a horizontal surface.

These data were integrated with local meteorological records which was obtained from *Visual Crossing* weather service (https://www.visualcrossing.com/) and include ambient air temperature (°C), precipitation type, weather conditions (Table1). Missing values in the combined dataset were imputed using linear interpolation for numeric variables, ensuring a complete time series without gaps. Categorical meteorological variables (e.g., precipitation type, conditions) were transformed into binary indicators via one-hot encoding [21], creating separate features for each unique category across the dataset. Features with near-zero variance were excluded to reduce redundancy, and all exogenous variables were normalized to a [0, 1] range using min-max scaling [22], based on training period to prevent data leakage. Table 1 summarises the key meteorological variables, their descriptions, types and unit.

Table 1: Meteorological Features.

Feature Name	Description	Type	Unit
GHI_W_m2	Global Horizontal Irradiance	Numeric	W/m²
temp	Ambient air temperature	Numeric	°C
preciptype	Type of precipitation (e.g., rain, none).	Categorical	N/A
conditions	Cloud conditions (e.g., clear, cloudy).	Categorical	N/A
icon	Weather icon descriptor (e.g., sunny, rainy)	Categorical	N/A
stations	Reporting weather station identifiers	Categorical	N/A
name	Location or station name	Categorical	N/A
dew	Dew point temperature	Numeric	°C.
humidity	Relative humidity	Numeric	%
precip	Precipitation amount	Numeric	mm
precipprob	Probability of precipitation	Numeric	%
windgust	Maximum wind gust speed	Numeric	km/h
windspeed	Mean wind speed	Numeric	km/h
winddir	Wind direction	Numeric	degrees
cloudcover	Cloud cover percentage	Numeric	%
solarradiation	Solar radiation	Numeric	W/m²
solarenergy	Solar energy accumulation	Numeric	MJ/m²
uvindex	Ultraviolet radiation index.	Numeric	N/A

Source: Authors, (2026).

To capture temporal and seasonal patterns inherent in load and PV dynamics, additional analytical features were engineered. These include basic components such as year, month number, day of the week, hour of the day, day of the year, and week of the year. Cyclic patterns were encoded using trigonometric transformations to represent periodicity continuously, avoiding discontinuities in numerical representations (Table 2). These exogenous features form the regressors for modelling, enabling the incorporation of weather-driven variability and temporal structures into net load simulations.

Table 2: Temporal Features.

Feature Name	Description	Type
year	Year	Numeric
month_num	Month of the year	Numeric
day_of_week_num	Day of the week,	Numeric
hour_of_day_num	Hour of the day	Numeric
day_of_year_num	Day of the year	Numeric
week_of_year_num	Week of the year	Numeric
hourofday_sin	Sine transformation for the hour of the day	Numeric
hourofday_cos	Cosine transformation for the hour of the day	Numeric
dayofweek_sin	Sine transformation for the day of the week	Numeric
dayofweek_cos	Cosine transformation for the day of the week	Numeric
month_sin	Sine transformation for the month of the year	Numeric
month_cos	Cosine transformation for the month of the year	Numeric
dayofyear_sin	Sine transformation for the day of the year	Numeric
dayofyear_cos	Cosine transformation for the day of the year	Numeric
period_wet	Binary indicator for the wet season	Binary
is_weekend	Binary indicator for weekends	Binary

Source: Authors, (2026).

II. Solar PV Generation Profile: The hourly GHI, is separated into Direct Normal Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI);

$$\mathrm{GHI}_t = \mathrm{DHI}_t + \mathrm{DNI}_t \cos\left(\theta_{z,t}\right) \tag{1}$$

Where the solar zenith angle $\theta_{z,t}$ is defined in terms of the geographical location latitude ϕ, the solar declination δ_t and the hour angle ω_t.

The direct and diffuse components are then mathematically projected from the horizontal plane onto the defined plane of the PV array. This array has a fixed tilt angle β and azimuth angle γ_p. The total Plane-of-Array (POA) irradiance, $G_{POA,t}$, on the tilted surface is the sum of the three components:

$$G_{\mathrm{POA},t} = G_{b,c,t} + G_{d,c,t} + G_{r,c,t} \tag{2}$$

Where $G_{b,c,t}$ is the direct beam irradiance incident on the collector plane at time t, $G_{d,c,t}$ is the diffuse irradiance on the collector plane and $G_{r,c,t}$ the ground-reflected irradiance on the collector plane. The calculated POA irradiance is converted into a realistic electrical power output. Meanwhile, a PV system is not perfectly efficient due to significant losses during the energy conversion process.

Losses arise from inverter inefficiency in converting DC to AC power, resistive losses in wiring, performance degradation at high temperatures, and energy reduction due to soiling. Therefore, these factors are aggregated into a single de-rating factor called the Performance Ratio (PR). For this model, a PR of 0.85 is used, signifying that 85% of the available light energy is successfully converted to usable AC power.

The power output is;

$$P_t = \frac{G_{\mathrm{POA},t}}{1000} \cdot \mathrm{PR} \cdot A \tag{3}$$

Where A is the panel area in squared meters and the division by 1000 converts W/m^2 to kW/m^2. This equation yields the final PV power generation profile in kilowatts, which serves as the foundational template for the subsequent net load simulation. Any resulting negative values (which can occur from model approximations during non-daylight hours) are set to zero to ensure physical realism. The solar PV generation profile was simulated for a fixed, non-tracking system where the substation is located. The model assumed a south-facing orientation with a shallow panel tilt of 7.38°. The R *solarR* package was used for the implementation [23]. The hourly PV profile serves as a standardised shape for PV generation which will be scaled to achieve a target penetration level at the level of each individual feeder. Let $L_i(t)$ be the hourly gross load on feeder i at hour t, and $P_{PV,\mathrm{raw}}(t)$ be the hourly solar PV system AC output from the standardised annual profile. The renewable penetration level is defined as the total annual energy generated by PV relative to the total annual energy consumed by the gross load for each specific feeder i. The total annual gross load energy for feeder i (denoted $E_{L,i}$) and the total annual energy from the raw PV profile shape (denoted $E_{PV,\mathrm{raw}}$) over the period of interest (one year) are given by:

$$E_{L,i} = \sum_{t=1}^{8760} L_i(t) \tag{4}$$

$$E_{PV,\mathrm{raw}} = \sum_{t=1}^{8760} P_{PV,\mathrm{raw}}(t) \tag{5}$$

The scaling factor λ_i that ensures the total scaled PV energy meets the target P_{target} penetration level for feeder i is given as:

$$\lambda_i = \frac{\left(\frac{P_{\mathrm{target}}}{100}\right) \times E_{L,i}}{E_{PV,\mathrm{raw}}} \tag{6}$$

Applying this feeder-specific scaling factor λ_i to the original hourly PV profile $P_{PV,\mathrm{raw}}(t)$ returns the scaled hourly PV generation $P_{PV,\mathrm{scaled},i}(t)$ for feeder i. This process is repeated for each of the three years (2022-2024). This preserves the diurnal and seasonal shape of the PV generation while adjusting its overall magnitude to meet the feeder's target annual energy contribution. To ensure generation is non-negative:

$$P_{PV,\mathrm{scaled},i}(t) = \max\left(0, \lambda_i \times P_{PV,\mathrm{raw}}(t)\right) \tag{7}$$

Therefore, the net load $P_{\mathrm{Net},i}(t)$ at feeder i with P_{target} energy penetration is given as:

$$P_{\mathrm{Net},i}(t) = L_i(t) - P_{PV,\mathrm{scaled},i}(t) \tag{8}$$

After calculating the net loads for each individual feeder, these feeder-level net loads $P_{\mathrm{Net},i}(t)$ are then aggregated upwards to their respective transformer and ultimately to the station level to form the complete net load hierarchical time series. This process ensures that the PV impact is simulated based on the specific load characteristics at the lowest level before being combined into higher-level net load series.

III. Forecast Model: Artificial neural networks (ANNs) represent a class of forecasting methods inspired by simplified mathematical models of the human brain, enabling the capture of complex nonlinear relationships between response variables and predictors [24]. At their core, ANNs consist of interconnected *neurons* organized into an input layer for predictors, an output layer for forecasts, and optional intermediate hidden layers. Simple ANNs without hidden layers are equivalent to linear regression models, where weights are optimized via learning algorithms to minimize cost functions such as mean squared error (MSE). The addition of hidden layers introduces nonlinearity through activation functions, such as the sigmoid, which combine weighted inputs and enhance robustness to outliers [25]. Weights are typically initialized randomly, constrained by a decay parameter to prevent overfitting, and refined through multiple training iterations, with results averaged for stability.

Hyperparameters, including the number of hidden layers and neurons, are often determined via cross-validation. For time series forecasting, NNAR models incorporate lagged values as inputs, analogous to linear autoregressive models [26], [27]. An NNAR(p,k) denotes p lagged inputs and k hidden neurons, while seasonal variants, NNAR$(p,P,k)_m$, include additional lags from prior seasons, m being the seasonal period of the time series. These models can replicate ARIMA structures without stationarity constraints, allowing for asymmetric cyclic patterns. The NNETAR() function from the *forecast* package is used for the implementation. The function automates lag selection for non-seasonal data or seasonally adjusted series. Forecasting proceeds iteratively, using historical data and prior predictions as inputs. External regressors are used, turning the model into a NNARX (NNAR with exogenous inputs).

IV. Forecast Reconciliation: A Hierarchical Time Series (HTS) is a multivariate time series with linear aggregation constraint [28]. Mathematically, an HTS can be represented as;

$$y_t = Sb_t \tag{9}$$

Where y_t is a vector comprising the *n* time series in all levels of the HTS at time *t*, b_t is a vector of the *m* most disaggregated series of the HTS and *S* is the summing matrix [29]. Given an HTS having a total of *n* time series out of which n_b are the bottom-level series, the reconciled forecast is given as;

$$\tilde{y}_h = SG_h\hat{y}_h \tag{10}$$

Where $\tilde{y}_h$ is the reconciled forecasts, $\hat{y}_h$ is the h-step ahead unreconciled (base) forecasts and G_h is an $n_b \times$ n matrix. The function of the G matrix is to map the unreconciled forecasts into the bottom level which are then summed up by *S* to yield the reconciled forecasts $\tilde{y}_h$. A constraint on *G* to make sure the unbiased base forecasts $\hat{y}$ results in an unbiased reconciled forecasts $\tilde{y}_h$ is given by $SGS = S$ [30]. The challenge of statistical forecast reconciliation is to finding the optimal G matrix. This can be achieved by single-level approach or by optimal forecast reconciliation approach. In the Bottom-Up (BU) approach, the forecast of the lowest level series is extracted by *G* and then aggregated by the summing matrix *S* to obtain forecast of the top-level time series.

For the BU, $G = \left[0_{n_h \times n_a}\ I_{n_h}\right]$, where *n* is the total number of series in the hierarchy, n_b is the number of the bottom-level series and the upper-level series are n_a = *n*-n_b, 0 and I are matrix of zeros and the identity matrix respectively. Building up the HTS from the most disaggregated level in BU method means that no information is lost. However, the bottom level series are usually noisy and poses forecasting accuracy challenges. Let V_h and W_h be the variance-covariance matrix of $\tilde{y}_h$ and $\hat{y}_h$with respect to the actual value $\boldsymbol{y}$ respectively. Then, $V_h = SGW_hG'S'$ and the *G* that minimises the error variance of $\tilde{y}_h$ is the one that minimises the trace of V_h which is found via optimisation to be $G = (S'W_h^{-1}S)^{-1}S'W_h^{-1}$[31]. Obtaining W_h is practically difficult and four simplifying assumptions have been presented in the literature. In the following section, $k_h > 0$;

I. Ordinary Least Squares (OLS) Reconciliation [30]: In this approach, $W_h = k_hI$, hence $G = (S'S)^{-1}S'$. This in effect makes G independent of the data and thus makes OLS reconciliation relatively faster to compute. However, the OLS method gives equal weights to all the constituent time series and does not consider inter-relationships within the HTS.

II. Weighted Least Square (WLS) Reconciliation [32]: Here, $W_h = k_h diag(\widehat{W})$ where $\widehat{W} = \frac{1}{T}\sum_{t=1}^{T} e_t\, e_t'$ and e_t is an n-dimensional vector of forecast residuals arranged in the same sequence as the data.

III. Structural Scaling (STR) Reconciliation [27]: Here, $W_h = k_h$diag$(S1)$where 1 is an *m*-dimensional unit vector. The combination weights therefore reflect the structure of the HTS as only the total source of error variance at each node is considered.

IV. Minimal Trace (MinT) [31]: Here $W_h = k_hW_1\boldsymbol{W_h} = \boldsymbol{k_hW_1}$, where $\boldsymbol{W_1}$ is the full one-step covariance matrix with the assumption that the covariance matrices are proportional to each other. Two ways of calculating $W_1$are by calculating the sample covariance (SAM) or to use a shrinkage estimator (SHR).

The R *FoReco* package was used in the implementation of the five reconciliation methods [33].

V. Addressing Negative Values in the Reconciled Forecasts: Net-load becomes negative once the power generated by the solar photovoltaic (PV) system exceeds the gross load. Since negative net-load is sometimes prohibited in non-battery-based systems (i.e., curtailment), the set-negative-to-zero (*sntz*) heuristic was applied to ensure the generated reconciled forecasts are non-negative [33]. The *sntz* heuristic is a simple post-processing step to ensure non-negative, coherent forecasts in a forecast hierarchy. It starts with reconciled forecasts from any standard method (which may include negatives) and adjusts only the bottom-level forecasts before re-aggregating via bottom-up. Let $\tilde{\mathbf{b}}_h$ be the reconciled bottom-level forecasts (potentially with negatives). The *sntz*-adjusted bottom-level forecasts are $\tilde{\mathbf{b}}_h^+ = \max(\tilde{\mathbf{b}}_h, \mathbf{0})$ and the final non-negative reconciled forecasts are $\tilde{\mathbf{y}}_h^+ = \mathbf{S}\tilde{\mathbf{b}}_h^+$. The method may however introduce slight bias compared to optimization-based methods, but it is computationally faster and often yields similar or better NRMSE in practice.

VI. Experimental Setup: Hierarchical gross Load, solar PV generation and net-load at 30% penetration level will be used in this investigation. The net-load forecasting and reconciliation strategy to be evaluated include,

A. Direct Approach (Fig. 2)

I. Form the historical hierarchical net load hierarchy from the historical hierarchical gross load and solar PV generation data.

II. Train one of the forecast models and generate hierarchical net-load forecast.

III. Apply a reconciliation method to this set of forecasts to enforce hierarchical consistency. The result is the first final forecast set.

B. Indirect Approach (Fig. 3)

I. Train one of the forecast models on the historical hierarchical gross load data and generate hierarchical gross load forecast.

II. Train the same forecast models on the historical hierarchical solar PV generation data and generate hierarchical solar PV generation forecast.

III. Subtract the generated hierarchical solar PV generation from the generated hierarchical gross load forecast to obtain the generated hierarchical net-load forecast

IV.Apply a reconciliation method to the hierarchical net-load forecast to enforce hierarchical consistency. The result is the first final forecast set.

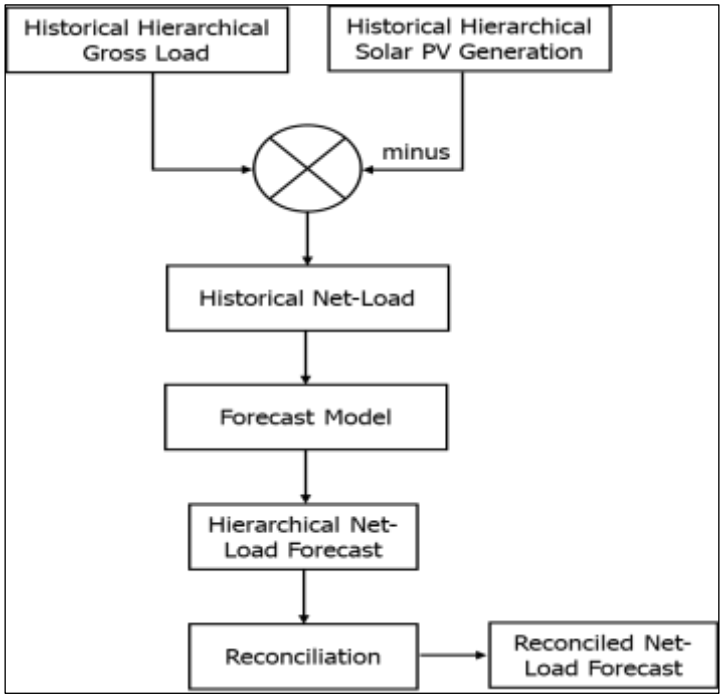

Figure 2: Direct Approach to HNLF.
Source: Authors, (2026).

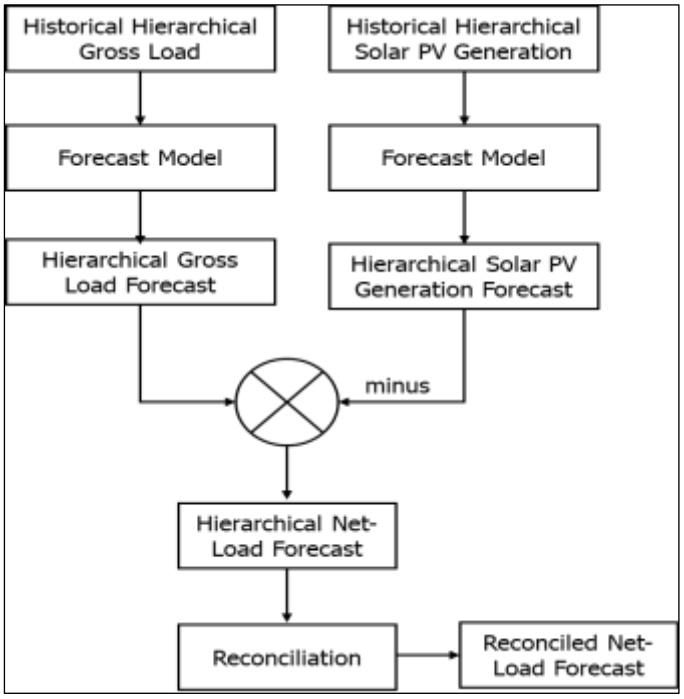

Figure 3: Indirect (Additive) Approach to HNLF.
Source: Authors, (2026).

VII. Evaluation: The final phase is the evaluation which employs a rolling forecast methodology (Fig 4), a technique designed to simulate real-world operational conditions. Instead of using a single, static train-test split, this approach iteratively creates multiple evaluation periods. In each iteration, the models are retrained on a fixed-size window of the most recent historical data (24 months) and then tested on the immediately following time period (one month). This entire process is repeated as the training and testing windows systematically slide forward through time. By aggregating the performance metrics from all iterations, this method provides a reliable assessment of a model's consistency and its true out-of-sample accuracy, ensuring that the results are not dependent on a single, arbitrary evaluation period. The hierarchical nature of the time series is an important factor while selecting metrics for evaluating forecast accuracy. Since series below aggregates to those above, the entire HTS will be on different scales, making absolute metrics not to be appropriate. Thus, this study adopts the Mean Absolute Scaled Error (MASE), Normalised Root Mean Square Error (NRMSE), Normalised Mean Bias Error (NMBE) which are scaled error metrics that enable comparison of the performances of forecasts across different aggregation levels.

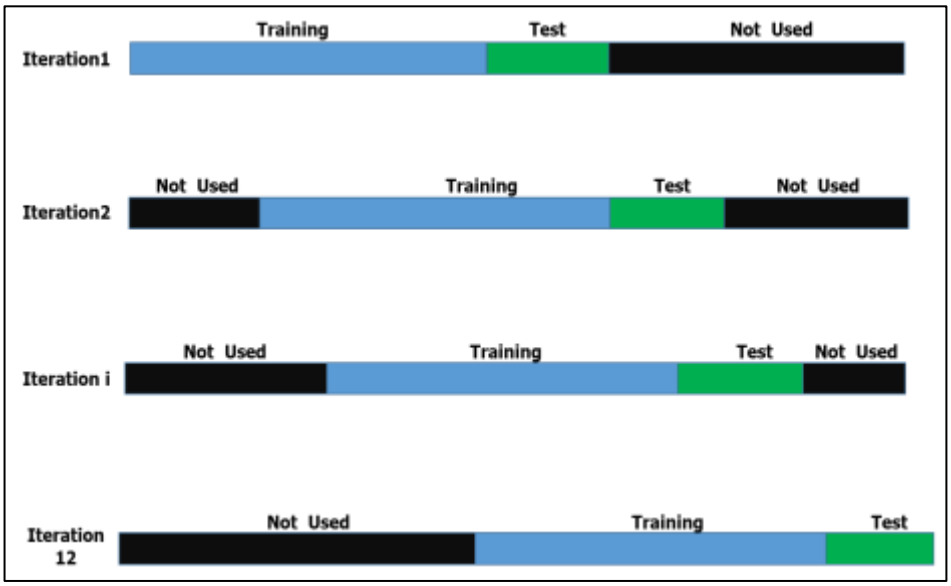

Figure 4: The Rolling Forecast Evaluation Strategy.
Source: Authors (2026).

These metrics have been used in reporting performance of HTS methods [33], [34], [35], [36]. Letting y_t be the actual values, $\hat{y}_t$ the forecasts, m the seasonal period, and n the length of the forecasts;

$$MASE = \frac{\frac{1}{n}\sum_{t=1}^{n} | y_t - \hat{y}_t |}{\frac{1}{T-m}\sum_{t=m+1}^{T} | y_t - y_{t-m} |} \quad (11)$$

Similarly,

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}}{\frac{1}{n}\sum_{t=1}^{n} y_t} \quad (12)$$

Also,

$$NMBE = \frac{\sum_{t=1}^{n} y_t - \hat{y}_t}{n \times \bar{y}} \quad (13)$$

Where T is the length of the in-sample data, $\bar{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$ is the mean of the actual series.When evaluating the performance of two forecasting models, statistical tests are essential to determine if one forecast is significantly more accurate than the other [37-39]. These tests typically analyse differences in loss functions, such as mean squared error (MSE) or absolute error, across a set of observations. The Wilcoxon signed-rank test is a non-parametric hypothesis test designed for comparing two related or paired samples to determine if there is a significant difference in their central tendencies [37]. Introduced as an alternative to the parametric paired t-test, it is particularly useful when the data do not satisfy normality assumptions or when the sample size is small. The null hypothesis is that the median of the differences is zero (no systematic difference in forecast accuracy), against the alternative that it is not. To determine if the strategies produced significantly different results, the Wilcoxon signed-rank test was applied. Rather than testing the final aggregated metrics, the test was performed on the paired distributions of the underlying, point-wise error components that constitute the metrics. Specifically, for MASE, the test compared the distributions of absolute errors (|actual - forecast|); for NRMSE, it compared the distributions of squared errors ((actual - forecast)²) and for NMBE, it compared the distributions of raw errors (actual - forecast).

IV. RESULTS AND DISCUSSIONS

I. Data Characteristics: Figure 5 shows the hourly profiles of the three key electrical quantities at the Station level for one full week (January 1st to 7th). The Net Load profile embodies the volatile characteristics of the gross load modulated by the solar PV profile. An indirect approach decomposes the net-load into components at the risk of summing their independent forecast errors, whereas the direct approach models the net-load, avoiding error compounding but tasking the model with learning the entangled component dynamics. The statistical profile of the grid's net load transforms significantly with increasing solar PV penetration (Table 3). While the mean and median net load predictably decrease across all hierarchical levels, this is contrasted by a sharp increase in volatility, as measured by the standard deviation. This heightened variability is a direct consequence of the *duck curve* phenomenon (Fig. 6), where deep midday load suppression creates a greater spread between the low solar-hours demand and the high evening demand [40]. This phenomenon appears to be more pronounced during the dry season than the wet season while being independent of day type (weekend or weekdays).]

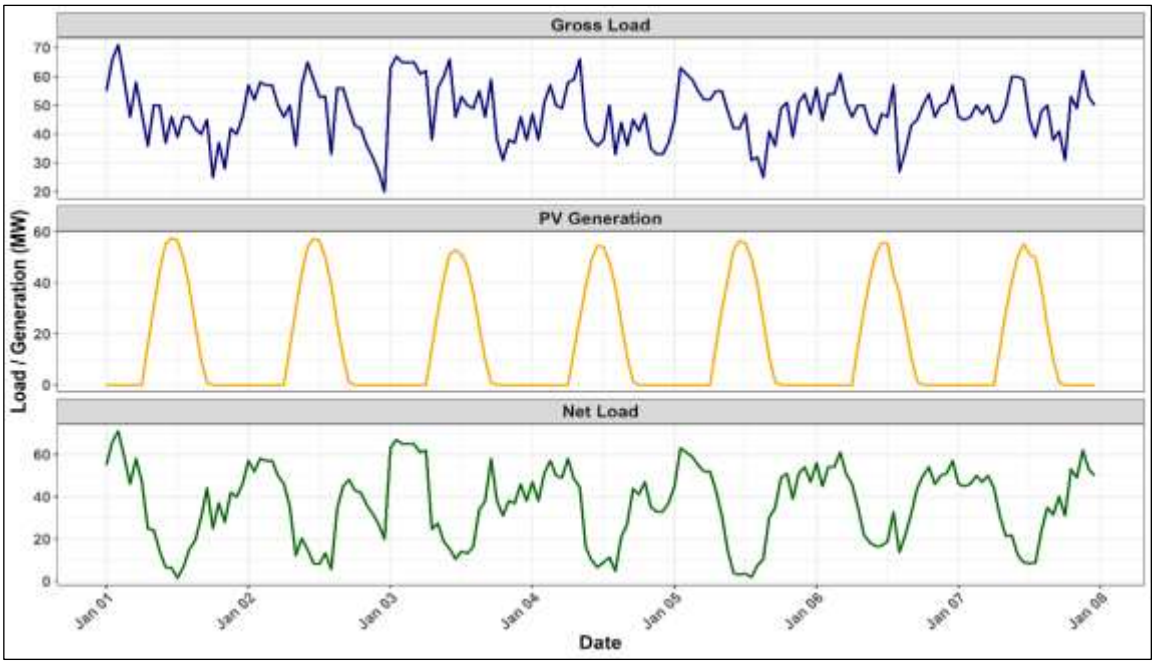

Figure 5: Hourly Profiles of the Three Key Electrical Quantities at The Station Level for A Week.
Source: Authors (2026).

Table 3: Statistical Summary of the Net-Load Data Across the Network Hierarchy.

Level	Penetration (%)	Mean (MW)	Std. Dev (MW)	Median (MW)	Min (MW)	Max (MW)
Level 0	0	46	14.7	47	0	103
	10	42.2	15	42.9	0	95.6
	20	38.6	16.8	39	0	92.3
	30	35.6	18.9	36.8	0	91.9
	40	33.2	20.7	35	0	91.5
	50	31.5	21.9	34	0	91.1
Level 1	0	15.3	6.6	15.7	0	37.7
	10	14.1	6.5	14.3	0	35
	20	12.9	6.9	13	0	34.7
	30	11.9	7.4	11.9	0	34.7
	40	11.1	7.8	11	0	34.7
	50	10.5	8.1	10.6	0	34.7
Level 2	0	5.1	3.2	5.5	0	18.3
	10	4.7	3.1	4.7	0	18.1
	20	4.3	3.1	4.3	0	17.9
	30	4	3.2	3.7	0	17.8
	40	3.7	3.2	3.4	0	17.7
	50	3.5	3.3	3.2	0	17.6

Source: Authors (2026).

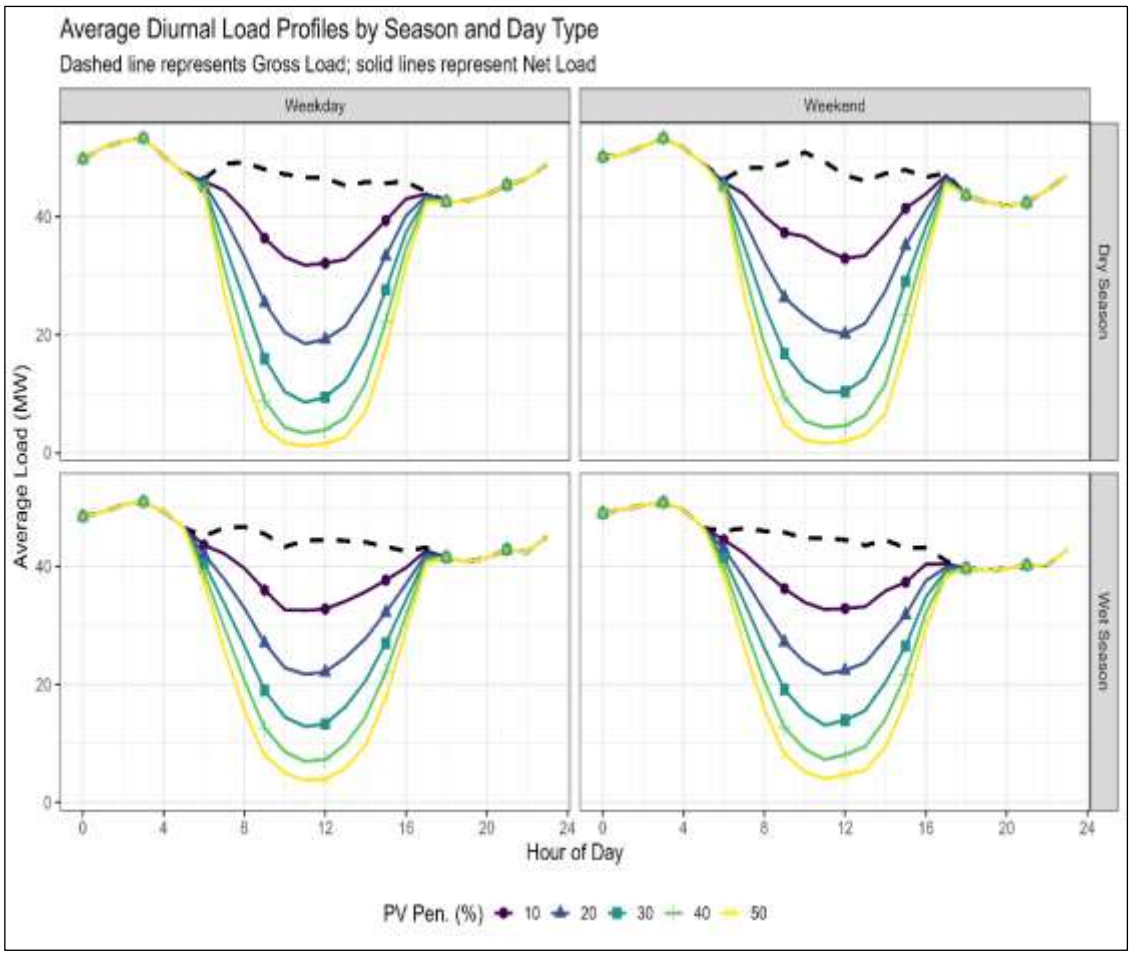

Figure 6: Average Diurnal Net-Load Profiles by Season and Day Type. The Dashed Line Represents the Original Gross Load.
Source: Authors (2026).

PV penetration also alters the predictability of net-load, as evident in the Autocorrelation Function (ACF) plot of Figure 7. The autocorrelation analysis of system net load under varying PV penetrations (0-50%) reveals a dominant 24-hour periodicity consistent with daily load cycles. At 0% PV, the ACF exhibits strong positive correlations at short lags with damped peaks at multiples of 24 hours, reflecting both short-term dependence and daily seasonality. As PV penetration increases, the diurnal cycle in the net load becomes more pronounced, with the ACF showing sharper oscillations and alternating positive (24-hour) and negative (12-hour) correlations.

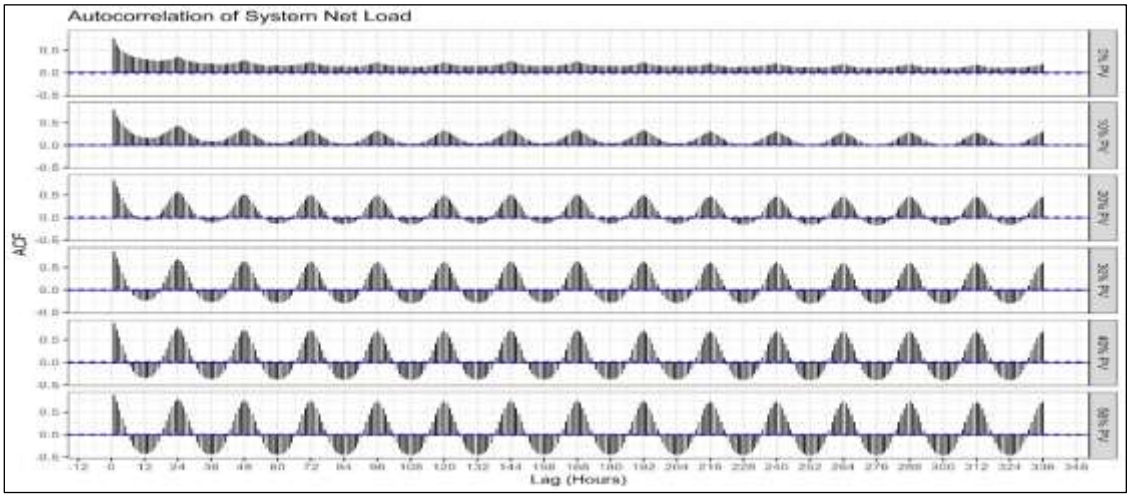

Figure 7: Autocorrelation of The System Net Load. The Correlation at The 24-Hour Lag (And Its Multiples) Is Visibly Weaker for Higher PV Penetration Levels.
Source: Authors(2026).

II. Performance of the Two Strategies: Figure 8 presents a visual analysis of the forecasting model applied to the Station Net Load. It compares a Direct forecasting strategy (Left panel, red colour), where models are trained on the Station Net Load series itself, against an Indirect strategy (Right panel, blue colour), where the Station forecast is derived from other component forecasts. The actual Station Net Load is shown in green for a single representative week. The actual net load exhibits a daily pattern in which the demand is lowest during the early morning hours and peaks in the evening. This cyclical behaviour is the primary feature that both forecasting models are attempting to capture. The Direct forecast successfully captures the overall timing and shape of the daily cycles. It correctly identifies when the peaks and troughs will occur. However, this model exhibits a bias related to the magnitude of the load. At almost every daily peak, the red forecast is noticeably higher than the green actual load. This indicates a tendency to overestimate demand during high-load periods. Conversely, during the lowest points of the day, the forecast is often just higher than the actual load, suggesting it also slightly under-estimates the lows.

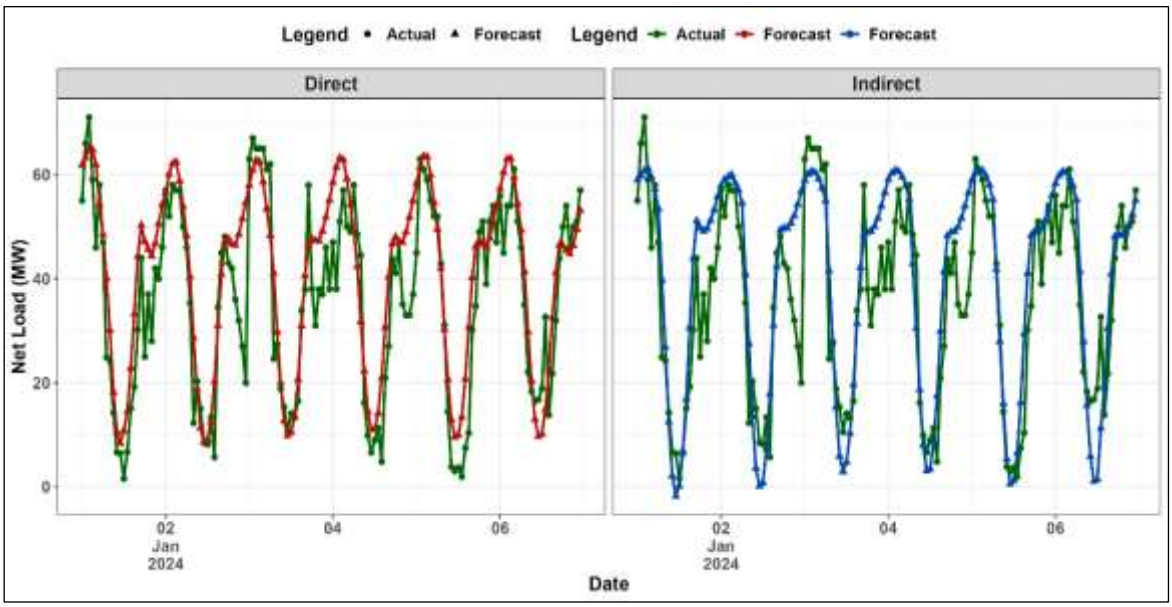

Figure 8: Time Series Plot of the Forecast of the Two Approaches and the Actual Values for the First Week of January 2024.
Source: Authors, (2026).

Like the Direct method, the Indirect forecast also does a good job of tracking the daily rhythm of the net load. However, this model also displays a consistent bias, but in the opposite direction of the Direct method. The Indirect forecast consistently fails to reach the height of the actual peaks. It underestimates demand during high-load periods. Its performance in the troughs appears often lower to the actuals compared to the Direct method. The Indirect model's forecast therefore appears smoother and more conservative. It is less responsive to the sharp peaks in the actual data, leading to its underestimation. The most important finding from this plot is that the two models have opposing systematic biases at the Station level. The Direct model over-forecasts, while the Indirect model under-forecasts. Neither model is consistently accurate in predicting the magnitude of peak load. The better model depends entirely on the business context and the cost associated with forecasting errors. If the cost of under-supplying power (e.g., needing to buy expensive power on the spot market) is very high, the Direct model's tendency to over-forecast might be preferable as a safer option.

If the cost of over-supplying power (e.g., wasted generation, selling excess at a loss) is the primary concern, the Indirect model's conservative, under-forecasting nature might be more desirable. Figure 9 displays boxplots which visualizes the distribution of forecast errors for the Direct and Indirect forecasting strategies. The Y-axis employs a symmetric logarithmic (symlog) scale, a technique necessary for visualizing distributions centred around zero that contain a wide range of values, including large outliers. This scale remains linear near the origin to allow for a clear comparison of the central tendencies, while compressing the tails to effectively display extreme errors without distortion. The Direct strategy exhibits a negative bias, as indicated by its median error being below the zero line. This demonstrates a systematic tendency to over-forecast the net load. The interquartile range (IQR), represented by the height of the box, is relatively compact which suggests that the middle 50% of its forecast errors are clustered within a tighter range, indicating a higher degree of consistency compared to the Indirect method.

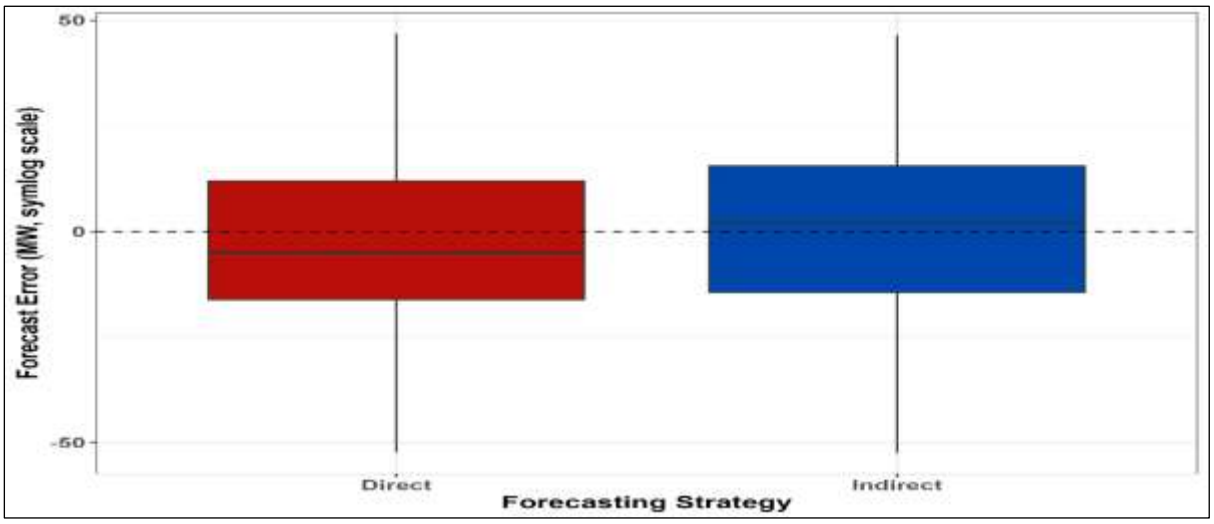

Figure 9: Forecast Error Distribution of the Two Strategies.
Source: Authors (2026).

The whiskers are also symmetrical, showing that the risk of large-magnitude errors is present for both over- and under-forecasting, though the central tendency remains biased towards over-forecasting. The error distribution of the Indirect strategy is more centred, with a median positioned slightly above the zero line. This signifies a minor positive bias, indicating a slight tendency to under-forecast the net load. This strategy displays higher forecast variance as the IQR is visibly larger than that of the Direct method, meaning its forecasts are less consistent, with the central 50% of errors covering a wider range of values. The error distribution is also asymmetrical with the upper whisker notably shorter than the lower one, highlighting a risk of large negative errors. This means the Indirect strategy is more prone to occasional, large-magnitude over-forecasting events. The plot reveals a distinct trade-off between the two forecasting strategies. The Direct method produces more consistent forecasts, but they are subject to systematic over-forecasting bias. In contrast, the Indirect method is less biased on average, but this comes at the cost of higher variance and an asymmetrical risk profile skewed towards major over-forecasting events.

The decision between the two would depend on the specific operational tolerance for predictable bias versus forecast inconsistency and the asymmetrical risk of large errors. The dumbbell plots in Figure 10-12 visualize the impact of different reconciliation methods on the accuracy and bias of the Direct and Indirect base forecasts. Each grey line connects the performance of an unreconciled *Base* forecast (grey point) to its *Reconciled* counterpart (coloured point), immediately revealing the magnitude and direction of the change. An improvement in accuracy is shown by a coloured point to the left of its grey partner, while an improvement in bias is shown by a coloured point closer to the vertical zero line. The MASE plot for both strategies showed that optimal reconciliation methods offer some degree of improvement in terms of accuracy over the base forecast while BU is the worst for both strategies. the Indirect strategy shows the most substantial improvements by the optimal reconciliation methods (Fig. 10). While the SHR method is the overall best for the Direct strategy, the STR method offered the biggest average improvements to the Indirect forecast.

These patterns are sustained in the NRMSE plot (Fig. 11), confirming the validity of the observations. The NMBE plot however tells a completely different story (Fig. 12). While the Indirect method again benefitted more from reconciliation, all reconciled forecast from the Direct strategy remain far from the zero line with reconciliation largely ineffective at correcting the large negative bias (over-forecasting) inherent in this approach. Conversely, for the Indirect strategy, BU is slightly effective at improving bias while optimal reconciliation methods largely worsening the forecast bias. These observations make it clear that the choice of forecasting strategy and reconciliation method must be driven by the primary operational objective. If the goal is maximum forecast accuracy, the plots provide evidence that the Direct strategy combined with SHR or OLS reconciliation is the superior choice. However, if the goal is to produce a less-biased forecast, the Indirect strategy combined with the Bottom-Up (BU) method is the only effective option.

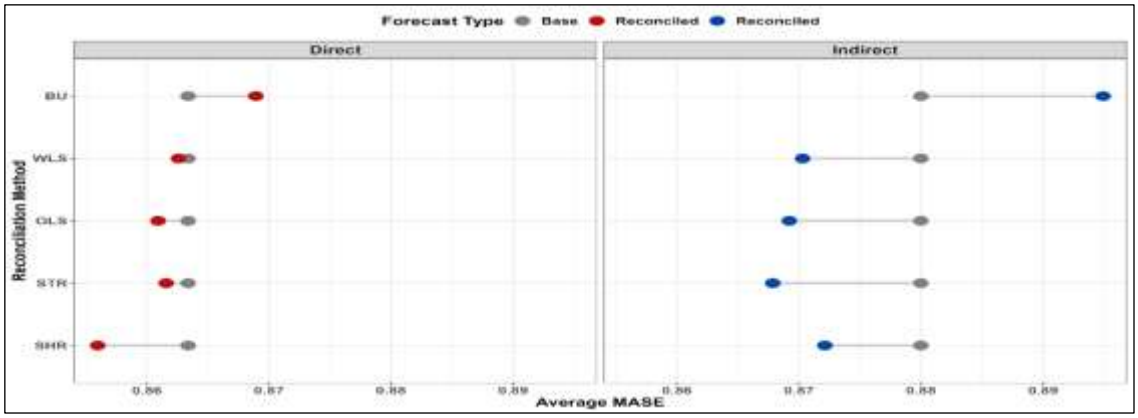

Figure 10: Dumbbell Plots of the MASE of the Direct and Indirect Forecasts.
Source: Authors (2026).

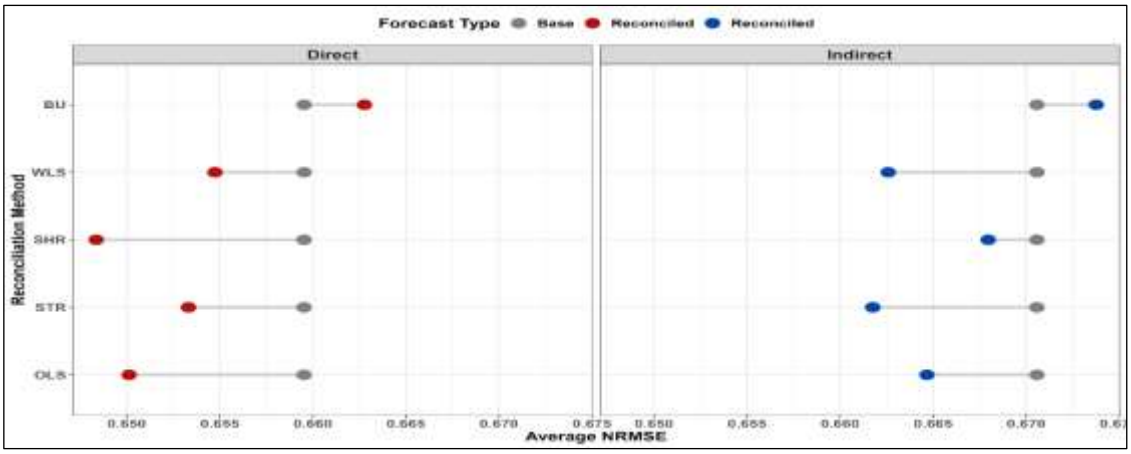

Figure 11: Dumbbell Plots of the NMBE of the Direct and Indirect Forecasts.
Source: Authors (2026).]

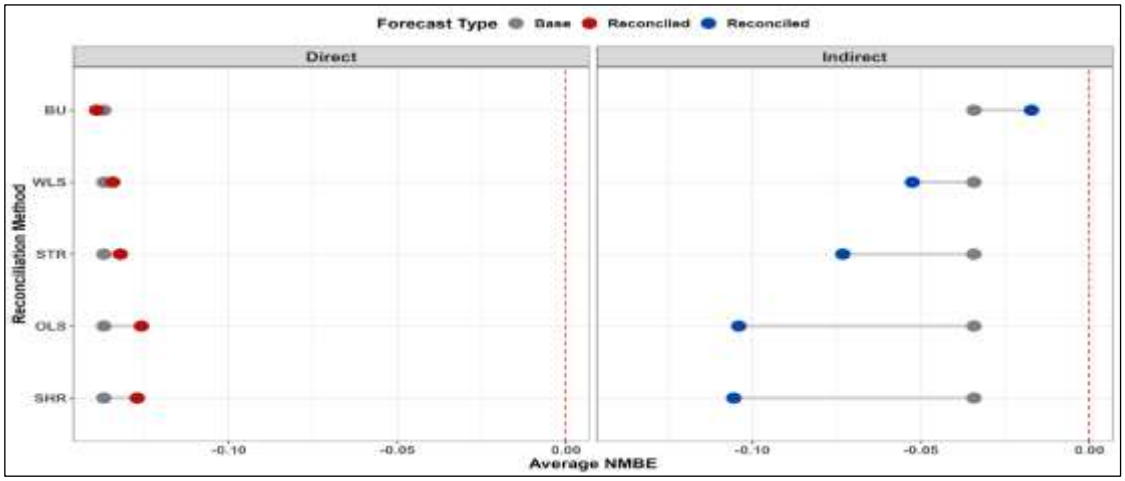

Figure 12: Dumbbell Plots of the NMBE of the Direct and Indirect Forecasts.
Source: Authors (2026).

Table 4 presents the results of a paired Wilcoxon signed-rank test to determine if there is a consistent, significant difference in performance between the Direct and Indirect forecasting strategies. The test is performed independently for each series at each Level of the hierarchy. The columns in the table are interpreted as follows:

I. Series: The series being evaluated.
II. Level: The hierarchical level at which the test was performed.
III. *p*_value: The probability of observing the performance difference between the two strategies if there were no actual underlying difference. A low *p*-value (less than 0.05) suggests that the observed difference is unlikely to be due to random chance.
IV. Median (Direct) / Median (Indirect): The median MASE value for each strategy. The median represents the typical performance, as it is the middle value of all error results and is not skewed by unusually large or small errors. Lower values indicate better typical performance.

V. Significance: A direct interpretation of the *p*-value. If the *p*-value is below the standard threshold of 0.05, the result is declared *Significant*, meaning there is statistical evidence of a real difference between the two strategies. Otherwise, it is *Not Significant*.
VI.Winner: This column identifies the strategy with the better (lower) median MASE, but only if the result is statistically significant. If the difference is not significant, no winner can be declared because the observed difference could be due to random chance.

Table 4: Results of a Paired Wilcoxon Signed-Rank Test to Compare the Two Strategies.

Series	Level	p_value	Median (Direct)	Median (Indirect)	Significance	Winner
Total	Station	1.27E-09	6.906	7.514	Significant	Direct
A	Transformers	1.15E-05	3.447	3.533	Significant	Direct
A1	Feeders	4.40E-87	2.362	2.848	Significant	Direct
A2	Feeders	2.08E-67	2.137	2.439	Significant	Direct
B	Transformers	2.08E-07	3.393	3.377	Significant	Indirect
B1	Feeders	4.13E-65	1.212	1.425	Significant	Direct
B2	Feeders	4.45E-32	0.511	0.599	Significant	Direct
B3	Feeders	1.49E-45	1.991	1.853	Significant	Indirect
B4	Feeders	1.0E-148	1.744	2.178	Significant	Direct
C	Transformers	1.83E-31	3.579	4.181	Significant	Direct
C1	Feeders	9.80E-04	2.020	2.204	Significant	Direct
C2	Feeders	1.40E-82	1.788	1.994	Significant	Direct
C3	Feeders	1.68E-03	1.358	1.456	Significant	Direct

Source: Authors (2026).

The result confirmed that the difference in performance between the two strategies is statistically significant for every single series analysed. All p-values are very small, indicating that the observed differences in accuracy are not due to random chance but reflect a genuine performance gap. With statistical significance established across the board, the *Winner* column clearly identifies the superior strategy based on the lower median absolute error. The Direct strategy is the overall winner, demonstrating superior performance at the highest aggregate level and for the vast majority of the individual components. At the Station level, The Direct method is significantly more accurate, with a median absolute error of 6.91 compared to the Indirect method's 7.51. The Direct strategy wins 2 out of the 3 transformer series (A and C) while the Indirect strategy shows a competitive edge for series B.

The dominance of the Direct strategy is most pronounced at the most disaggregated level. It is the statistically significant winner for 8 out of the 9 feeders. The Indirect method was only superior for the B3 feeder. The statistical analysis provides a definitive conclusion that the Direct forecasting strategy is demonstrably and significantly more accurate than the Indirect strategy for this system**.** Its superior performance is a consistent and statistically validated characteristic observed at the total system level and across almost all of its constituent parts. While the Indirect method may perform better in isolated cases, the evidence overwhelmingly supports the selection of the Direct strategy as the more reliable and accurate base forecasting approach.

V. CONCLUSIONS

The study provides an empirical answer for the direct versus indirect netload forecasting debate with insights from an hierarchical time series point of view, unlike prior studies which were often limited to a single aggregation level and reported conflicting outcomes [6], [11], [41]. This work demonstrates that for the network studied, the Direct approach yields more accurate base forecasts, potentially because a single model can better capture the non-linear interactions between load and solar PV generation that separate additive models may miss and also avoided the compounding of errors from the two component forecasts. The study also quantified the accuracy-bias trade-off inherent in the choice of reconciliation method, providing a practical framework for operational decision-making.

This research empirically demonstrates that while optimal reconciliation methods are designed to improve coherence and minimise error variance [31], they operate by preserving the unbiasedness of the base forecasts and thus cannot fix an already biased model [30]. The practical implication for grid operators is an evidence-based choice. For strategic planning where minimizing error magnitude is paramount, the Direct strategy combined with SHR reconciliation is demonstrably superior. However, for real-time operational scheduling where an unbiased forecast is critical to prevent systemic over-commitment of generation resources, the Indirect strategy with BU reconciliation presents the most viable low-bias alternative. This directly addresses the need to align forecast evaluation metrics with the specific operational decisions being made [42].

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Oyebola Olasupo, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Methodology: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Investigation: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Discussion of results: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Writing – Original Draft: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Writing – Review and Editing: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Resources: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Supervision: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.
Approval of the final text: Oyebola Olasupo*, Temitope Raphael Ayodele, Ayodeji Samson Olatunji Ogunjuyigbe.

I.REFERENCES

[1] A. Azeem, I. Ismail, S. M. Jameel, and V. R. Harindran, “Electrical Load Forecasting Models for Different Generation Modalities: A Review,” IEEE Access, vol. 9, pp. 142239–142263, 2021, doi: 10.1109/ACCESS.2021.3120731.

[2] A. Perçuku, D. Minkovska, and L. Stoyanova, “Big Data And Time Series Use In Short Term Load Forecasting In Power Transmission System,” Procedia Comput. Sci., vol. 141, pp. 167–174, 2018, doi: 10.1016/j.procs.2018.10.163.

[3] S. I. Kampezidou and S. Grijalva, “Distribution transformers short-term load forecasting models,” in IEEE Power and Energy Society General Meeting, IEEE Computer Society, Nov. 2016. doi: 10.1109/PESGM.2016.7741174.

[4] Y. Chu, H. T. C. Pedro, A. Kaur, J. Kleissl, and C. F. M. Coimbra, “Net load forecasts for solar-integrated operational grid feeders,” Sol. Energy, vol. 158, pp. 236–246, Dec. 2017, doi: 10.1016/j.solener.2017.09.052.

[5] S. E. Razavi, A. Arefi, G. Ledwich, G. Nourbakhsh, D. B. Smith, and M. Minakshi, “From Load to Net Energy Forecasting: Short-Term Residential Forecasting for the Blend of Load and PV Behind the Meter,” IEEE Access, vol. 8, pp. 224343–224353, 2020, doi: 10.1109/ACCESS.2020.3044307.

[6] G. Aburiyana, H. Aly, and T. Little, “Direct Net Load Forecasting Using Adaptive Neuro Fuzzy Inference System,” in 2021 IEEE Electrical Power and Energy Conference (EPEC), Toronto, ON, Canada: IEEE, Oct. 2021, pp. 131–136. doi: 10.1109/EPEC52095.2021.9621457.

[7] X. Sun and C. Jin, “Impacts of Solar Penetration on Short-Term Net Load Forecasting at the Distribution Level,” in 2021 IEEE 4th International Electrical and Energy Conference (CIEEC), Wuhan, China: IEEE, May 2021, pp. 1–6. doi: 10.1109/CIEEC50170.2021.9510770.

[8] S. Kerkau, S. Sepasi, H. O. R. Howlader, and L. Roose, “Day-Ahead Net Load Forecasting for Renewable Integrated Buildings Using XGBoost,” Energies, vol. 18, no. 6, p. 1518, Mar. 2025, doi: 10.3390/en18061518.

[9]S. Sreekumar, K. C. Sharma, and R. Bhakar, “Grey System Theory Based Net Load Forecasting for High Renewable Penetrated Power Systems,” Technol. Econ. Smart Grids Sustain. Energy, vol. 5, no. 1, p. 21, Dec. 2020, doi: 10.1007/s40866-020-00094-4.

[10] P. Kobylinski, M. Wierzbowski, and K. Piotrowski, “High-resolution net load forecasting for micro-neighbourhoods with high penetration of renewable energy sources,” Int. J. Electr. Power Energy Syst., vol. 117, p. 105635, May 2020, doi: 10.1016/j.ijepes.2019.105635.

[11] M. Alipour, J. Aghaei, M. Norouzi, T. Niknam, S. Hashemi, and M. Lehtonen, “A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration,” Energy, vol. 205, p. 118106, Aug. 2020, doi: 10.1016/j.energy.2020.118106.

[12] O. Rubasinghe et al., “Highly accurate peak and valley prediction short-term net load forecasting approach based on decomposition for power systems with high PV penetration,” Appl. Energy, vol. 333, p. 120641, Mar. 2023, doi: 10.1016/j.apenergy.2023.120641.

[13] J. Silva-Rodriguez, E. Raffoul, and X. Li, “LSTM-Based Net Load Forecasting for Wind and Solar Power-Equipped Microgrids,” in 2024 56th North American Power Symposium (NAPS), El Paso, TX, USA: IEEE, Oct. 2024, pp. 1–6. doi: 10.1109/NAPS61145.2024.10741801.

[14] A. Stratman, T. Hong, M. Yi, and D. Zhao, “Net Load Forecasting with Disaggregated Behind-the-Meter PV Generation,” in 2022 IEEE Industry Applications Society Annual Meeting (IAS), Detroit, MI, USA: IEEE, Oct. 2022, pp. 1–5. doi: 10.1109/IAS54023.2022.9940025.

[15] C. Thepprom, N. Nupairoj, and P. Vateekul, “A Deep Learning Framework for Net Load Forecasting With Unsupervised Behind-the-Meter Disaggregated Data,” IEEE Access, vol. 12, pp. 94958–94971, 2024, doi: 10.1109/ACCESS.2024.3425590.

[16] A. Falces et al., “Short-term net load forecast in distribution networks with PV penetration behind the meter,” Energy Rep., vol. 9, pp. 115–122, May 2023, doi: 10.1016/j.egyr.2022.12.103.

[17] A. Kaur, L. Nonnenmacher, and C. F. M. Coimbra, “Net load forecasting for high renewable energy penetration grids,” Energy, vol. 114, pp. 1073–1084, Nov. 2016, doi: 10.1016/j.energy.2016.08.067.

[18] G. Tziolis et al., “Short-term electric net load forecasting for solar-integrated distribution systems based on Bayesian neural networks and statistical post-processing,” Energy, vol. 271, p. 127018, May 2023, doi: 10.1016/j.energy.2023.127018.

[19] M. Beichter, K. Phipps, M. M. Frysztacki, R. Mikut, V. Hagenmeyer, and N. Ludwig, “Net load forecasting using different aggregation levels,” Energy Inform., vol. 5, no. S1, p. 19, Sep. 2022, doi: 10.1186/s42162-022-00213-8.

[20] N. Sharma, R. Bhakar, and P. Jain, “Optimal reconciliation of hierarchical wind power forecasts of correlated wind farms,” Sustain. Energy Grids Netw., vol. 35, p. 101091, Sep. 2023, doi: 10.1016/j.segan.2023.101091.

[21] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, “Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network,” IEEE Trans. Smart Grid, vol. 10, no. 1, pp. 841–851, Jan. 2019, doi: 10.1109/TSG.2017.2753802.

[22] S. Bhanja and A. Das, “Impact of Data Normalization on Deep Neural Network for Time Series Forecasting,” Jan. 07, 2019, arXiv: arXiv:1812.05519. Accessed: Dec. 02, 2024. [Online]. Available: http://arxiv.org/abs/1812.05519

[23] O. Perpiñán, “solaR : Solar Radiation and Photovoltaic Systems with R,” J. Stat. Softw., vol. 50, no. 9, 2012, doi: 10.18637/jss.v050.i09.

[24] I. N. da Silva, D. Hernane Spatti, R. Andrade Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, Artificial Neural Networks. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-43162-8.

[25] A. D. Rasamoelina, F. Adjailia, and P. Sincak, “A Review of Activation Function for Artificial Neural Network,” in 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI), IEEE, Jan. 2020, pp. 281–286. doi: 10.1109/SAMI48414.2020.9108717.

[26] H. Qiao, K. Chalermyanont, and R. Duangsoithong, “Hour-Ahead Power Load Demand Time Series Forecasting Using Four Methods in Three Cases,” in 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Pattaya, Thailand: IEEE, Jul. 2019, pp. 593–596. doi: 10.1109/ECTI-CON47248.2019.8955222.

[27] R. J. Hyndman and G. Athanasopoulos, Forecasting: principles and practice. Melbourne: OTexts, 2021.

[28] J. P. Karmy and S. Maldonado, "Hierarchical time series forecasting via Support Vector Regression in the European Travel Retail Industry," Expert Syst. Appl., vol. 137, pp. 59–73, Dec. 2019, doi: 10.1016/j.eswa.2019.06.060.

[29] G. Athanasopoulos, R. J. Hyndman, N. Kourentzes, and A. Panagiotelis, "Forecast reconciliation: A review," Int. J. Forecast., vol. 40, no. 2, pp. 430–456, Apr. 2024, doi: 10.1016/j.ijforecast.2023.10.010.

[30] R. J. Hyndman, R. A. Ahmed, G. Athanasopoulos, and H. L. Shang, "Optimal combination forecasts for hierarchical time series," Comput. Stat. Data Anal., vol. 55, no. 9, pp. 2579–2589, Sep. 2011, doi: 10.1016/j.csda.2011.03.006.

[31] S. L. Wickramasuriya, G. Athanasopoulos, and R. J. Hyndman, "Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization," J. Am. Stat. Assoc., vol. 114, no. 526, pp. 804–819, Apr. 2019, doi: 10.1080/01621459.2018.1448825.

[32] R. J. Hyndman, A. J. Lee, and E. Wang, "Fast computation of reconciled forecasts for hierarchical and grouped time series," Comput. Stat. Data Anal., vol. 97, pp. 16–32, May 2016, doi: 10.1016/j.csda.2015.11.007.

[33] T. Di Fonzo and D. Girolimetto, "Spatio-temporal reconciliation of solar forecasts," Sol. Energy, vol. 251, pp. 13–29, Feb. 2023, doi: 10.1016/j.solener.2023.01.003.

[34] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," 2005. [Online]. Available: http://www.forecasters.org/data/m3comp/m3comp.htm

[35] D. Yang, H. Quan, V. R. Disfani, and C. D. Rodríguez-Gallegos, "Reconciling solar forecasts: Temporal hierarchy," Sol. Energy, vol. 158, pp. 332–346, Dec. 2017, doi: 10.1016/j.solener.2017.09.055.

[36] G. M. Yagli, D. Yang, and D. Srinivasan, "Reconciling solar forecasts: Sequential reconciliation," Sol. Energy, vol. 179, pp. 391–397, Feb. 2019, doi: 10.1016/j.solener.2018.12.075.

[37] B. Rosner, R. J. Glynn, and M. T. Lee, "The Wilcoxon Signed Rank Test for Paired Comparisons of Clustered Data," Biometrics, vol. 62, no. 1, pp. 185–192, Mar. 2006, doi: 10.1111/j.1541-0420.2005.00389.x.

[38] G. Dudek, "Short-term load forecasting using Theta method," in E3S Web of Conferences, EDP Sciences, Feb. 2019. doi: 10.1051/e3sconf/20198401004.

[39] Z. Zhang, W.-C. Hong, and J. Li, "Electric Load Forecasting by Hybrid Self-Recurrent Support Vector Regression Model With Variational Mode Decomposition and Improved Cuckoo Search Algorithm," IEEE Access, vol. 8, pp. 14642–14658, 2020, doi: 10.1109/ACCESS.2020.2966712.

[40] I. Calero, C. A. Canizares, K. Bhattacharya, and R. Baldick, "Duck-Curve Mitigation in Power Grids With High Penetration of PV Generation," IEEE Trans. Smart Grid, vol. 13, no. 1, pp. 314–329, Jan. 2022, doi: 10.1109/TSG.2021.3122398.

[41] G. Tziolis, A. Livera, A. Michail, G. Makrides, and G. E. Georghiou, "Direct Against Indirect Short-Term Net Load Forecasting Using Machine Learning Principles for Renewable Microgrids," in 2023 IEEE International Smart Cities Conference (ISC2), Bucharest, Romania: IEEE, Sep. 2023, pp. 1–5. doi: 10.1109/ISC257844.2023.10293666.

[42] R. Hollyman, F. Petropoulos, and M. E. Tipping, "Understanding forecast reconciliation," Eur. J. Oper. Res., vol. 294, no. 1, pp. 149–160, Oct. 2021, doi: 10.1016/j.ejor.2021.01.017.