Extending Weather Normalization to Time-Series Feeder-Level Forecasting for Distribution Planning

Lucas Almeida
Electric Power Research Institute
Palo Alto, CA, USA
lalmeida@epri.com

Jouni Peppanen
Electric Power Research Institute
Palo Alto, CA, USA
jpeppanen@epri.com

***Abstract*—There is increasing interest to extend feeder-level forecasts from conventional peak-load focused approaches to time-series 8,760-hour methodologies due to electrification, distributed energy resources (DERs), and climate-driven variability. This paper examines alternative weather-normalization methods for time-series load forecasting at the distribution level and evaluates practical weather-variable designs within an additive regression framework. Four normalization methods—calendar-day averaging, rank-and-average, Typical Meteorological Year (TMY), and multi-year simulation with quantiles—are compared conceptually and through a utility case study. Results indicate that calendar and rank-and-average approaches yield internally consistent 8,760 profiles suitable for year-over-year comparisons, while quantile-based multi-year simulation supports risk-aware planning; TMY can diverge from other methods for feeder-level studies. Feature selection also matters: temperature–humidity blends with short persistence tend to improve energy accuracy, whereas degree-hour terms favor peak accuracy. Guidance is offered on matching method and features to planning objectives.**

***Index Terms*— Distribution Planning, Electrification, Load Forecasting, Weather Normalization**

I. Introduction

Load forecasting is fundamental to distribution planning. Traditional approaches emphasized annual energy and peak demand, which were adequate for systems dominated by centralized generation and relatively stable demand. Increasing electrification, distributed energy resources (DERs), data centers, and climate-driven variability now introduce greater uncertainty and localization, motivating a shift toward feeder-level 8,760-hour time-series forecasting to capture load dynamics and geographically heterogeneous adoption patterns.

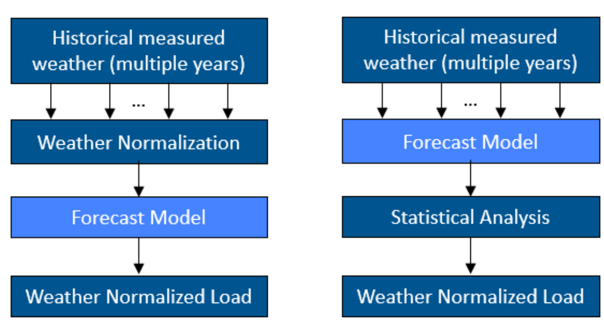
A central challenge in this paradigm is weather normalization, which adjusts load profiles to representative or scenario-based weather conditions to enable consistent comparison and defensible planning decisions. Conventional peak-based normalization methods are insufficient for time-series applications, where hourly weather impacts must be represented across the full year. Prior work shows that explicit weather constructs improve forecast accuracy and interpretability [1], and industry guidance increasingly emphasizes normalization for risk-based and probabilistic planning, particularly under growing climate variability [2], [3].

This work is part of the Load Forecasting Initiative (LFI) [4] led by the Electric Power Research Institute (EPRI). The paper focuses on developing practical guidance for weather normalization in feeder-level 8,760-hour load forecasting to support mid-term and long-term distribution planning. Building on recent EPRI research and industry surveys [5], [6], [7], [8], [9], it examines normalization challenges driven by electrification and climate variability through a representative feeder case study and a comparative evaluation of alternative methods. The broader LFI research by EPRI also explores forecast allocation, alignment between transmission and distribution planning, among numerous other forecasting areas, which are outside the scope of this paper.

This paper synthesizes four weather normalization methods for annual distribution planning load forecasting, applies engineered weather variables that incorporate temperature-humidity blends and short-term persistence, and evaluates trade-offs between energy accuracy and peak prediction under different modeling and normalization strategies.

II. Weather Normalization Background

Weather normalization adjusts historical load data to remove atypical weather effects and reveal underlying demand trends, enabling consistent comparison and defensible planning decisions across applications such as distribution planning, DER integration, and rate design [1], [2], [10], [11]. While traditional normalization methods were developed for annual energy and peak forecasting, extending them to feeder-level 8,760-hour applications requires explicit representation of nonlinear weather–load relationships using engineered inputs that reflect temperature, humidity, and temporal interactions [1], [10], [12], [13]. Accordingly, normalization approaches can be broadly categorized as historical-based methods, which apply predefined typical weather conditions (e.g., calendar averaging, rank-based averaging, TMY), and probabilistic simulation methods, which generate multiple weather scenarios to characterize variability and risk (Figure 1) [10], [11], [13].



(a) Historical weather data (b) Probabilistic simulations
Figure 1. Weather normalization approaches

979-8-3315-8138-1

III. Proposed Methodology

This section outlines the approach applied in this paper for evaluating weather normalization for distribution-level load forecasting. It begins by describing the case study context and data sources, including feeder-level load histories and associated weather records. Next, it details Distribution Planning Load Forecasting Framework developed by EPRI, emphasizing the selection of weather-sensitive features and error metrics used for performance evaluation. Finally, it describes the weather normalization methods used, noting their conceptual differences and impact on planning.

A. Case Study

The evaluation uses feeder-level hourly load data from an U.S. utility serving approximately half a million customers with a system peak in the order of 2 GW in 2022. The selected case feeder has a peak demand of approximately 12 MW and serves about 4,000 customers, representing a winter-peaking circuit with minimal distributed generation penetration. This feeder was chosen to illustrate weather normalization challenges under conditions of weather-driven load behavior and seasonal variability. The selected feeder is representative of many winter-peaking residential circuits where weather sensitivity dominates load behavior and DER penetration is currently low. While the presented results are not intended to be statistically universal, this controlled case enables clear isolation of normalization effects without confounding factors.

Historical load data spans four consecutive years (2016–2019), providing sufficient coverage to train the model. Hourly weather data were sourced from Open-Meteo (ERA5-based temperature and humidity) and NSRDB (MERRA2-based solar and meteorological variables), including Typical Meteorological Year (TMY) sequences for comparative analysis. Distributed generation and electrification impacts are not modeled explicitly due to low penetration in the selected feeder, but the discussed load forecasting methodology is extensible to scenarios with higher DER adoption.

B. EPRI Distribution Planning Load Forecasting Framework

The forecasting framework (Figure 2) is structured as a multi-step process to ensure consistency across system level and feeder granularity. The forecasting process begins with scenario definition, ensuring assumptions for electrification, DER adoption, and economic growth are aligned across system and distribution levels. Next, data collection and preparation integrate SCADA load histories, GIS feeder topology, weather data, and customer information, followed by preprocessing for quality and consistency. Model training applies time-series decomposition with engineered weather regressors to capture seasonal, trend, and temperature-humidity effects. A baseline forecast represents organic feeder growth under normalized weather, while load modifiers incorporate spot loads, electrification, and DER impacts. Finally, aggregation and alignment reconcile bottom-up feeder forecasts with top-down system projections to maintain consistency across planning horizons.

This paper focuses on the steps most critical to weather normalization: input data preparation, model training and parameter optimization, and baseline forecast generation. For this case study, input preparation includes time-series 8,760-hour historical load data spanning four consecutive years (2016–2019) and hourly weather data. While out of the scope of this paper, weather impacts should also be considered when aligning system-level and distribution-level forecasts and allocating system-level forecast information to distribution level.

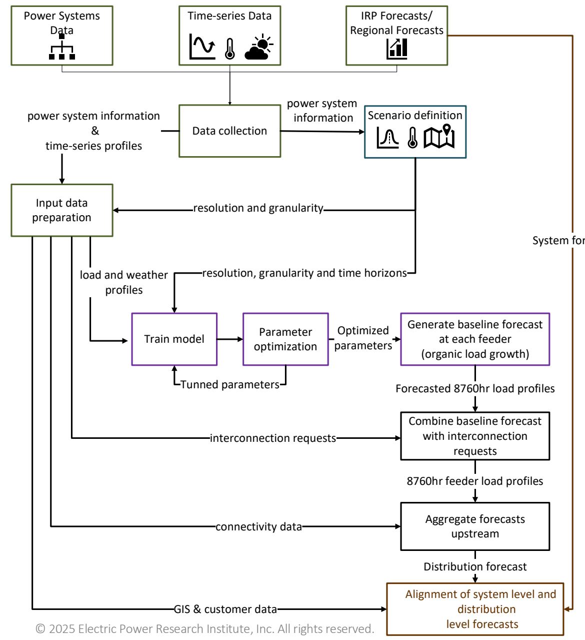


Figure 2. EPRI Distribution Planning Load Forecasting Framework

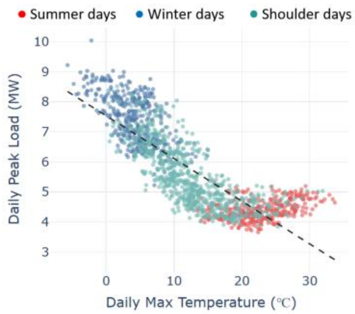
1) Weather Variable

Weather variables for load forecasting can be constructed in multiple ways, typically as functions of ambient temperature, humidity, wind speed, and cloud cover. Common formulations include dry bulb temperature, wet bulb temperature, dew point, apparent temperature, temperature-humidity, and degree-day metrics. The choice of formulation and transformation method significantly influences how well the variable captures nonlinear weather–load relationships, impacting both model accuracy and regional applicability. This study adopts a methodology based on EPRI Report [14], using an hourly temperature variable (TV) as the average of wet bulb and dry bulb temperatures, as shown in (1).

$$TV_t = \frac{DB_t + WB_t}{2} \quad (1)$$

Aggregated feeder loads respond differently to weather depending on customer composition. Evaluating this sensitivity is essential as inappropriately applying a temperature-driven model to a non-weather-sensitive feeder can lead to poor accuracy. For the study feeder, Figure 3 shows daily peak load plotted against the daily maximum temperature variable. The scatter plot reveals a clear winter-peaking pattern: lower temperatures correspond to higher loads, indicating electric heating influence. This strong correlation confirms that weather normalization is critical for this feeder, as accurate modeling of temperature effects improves both baseline forecasts and planning reliability.

979-8-3315-8138-1



© 2025 Electric Power Research Institute, Inc. All rights reserved.
Figure 3. Daily peak load against daily maximum TV

2) *Forecast Model*

Many modeling approaches can support distribution planning forecasts. For this case study, an additive time-series decomposition model with piecewise linear regression trends is applied, representing feeder load as the sum of long-term trend, deterministic seasonal components, and weather-driven effects. Seasonal terms capture recurring intra-day, weekly, and annual patterns, while the piecewise linear trend allows gradual slope changes to reflect structural evolution in feeder demand. Weather impacts enter the model additively through engineered temperature-based regressors that capture both instantaneous and persistent effects. Model variants discussed in this paper differ only in the selection and aggregation of these weather inputs, and candidate configurations are evaluated through cross-validated testing across historical years using performance metrics relevant to distribution planning, including overall profile accuracy, annual energy deviation, and seasonal peak error.

Weather variables are explicitly engineered to reflect local sensitivity. To ensure consistent terminology, this paper uses degree-hour (DH) to denote hourly heating and cooling load drivers. Accordingly, HDH and CDH represent hourly heating and cooling degree hours, while HDH24, CDH24, HDH48, and CDH48 denote 24-hour and 48-hour accumulated degree-hour terms. The adopted methodology begins with the hourly temperature variable (TV) mentioned before. To capture thermal inertia and nonlinear effects, additional transformations include: a 3-hour rolling TV (R3) for short-term smoothing, hourly degree-hour terms (HDH and CDH), and 24-hour and 48-hour accumulated degree hours (HDH24, CDH24, HDH48, and CDH48) to capture sustained heating and cooling effects.

The model was trained using four years of hourly load data (2016–2019). To ensure independence from training conditions and avoid atypical operational patterns, 2022 was reserved as an out-of-sample validation year. Intermediate years were excluded from validation to maintain a single, consistent benchmark year for all model and normalization comparisons. As presented in Table 1, eleven variants were tested, each incorporating different combinations of weather variables to identify the most accurate configuration through cross-validation. Performance was assessed using Mean Absolute Percentage Error (MAPE) for hourly profiles, annual energy deviation (APE-Energy), and seasonal/annual peak deviation (APE-Peak). This approach ensures that the selected regressor set aligns with feeder-specific weather sensitivity and planning objectives.

TABLE 1. FORECAST MODELS AND EXPLANATORY VARIABLES

Forecast Model	Variables				
	TV	*R3*	*HDH & CDH*	*HDH24 & CDH24*	*HDH48 & CDH48*
1	✓				
2		✓			
3	✓	✓			
4			✓		
5				✓	
6					✓
7	✓		✓		
8		✓		✓	
9		✓			✓
10	✓	✓	✓		
11	✓	✓	✓	✓	✓

© 2025 Electric Power Research Institute, Inc. All rights reserved.

3) *Weather Normalization*

After selecting the weather variable, assessing feeder sensitivity, and constructing the forecast model, the final step applies four weather normalization methods to generate 16 years of time-series load forecasts (2020–2035):

- **TMY**: Represents a composite year built from historical averages, repeated across the forecast horizon.
- **Calendar-Day Averaging**: Computes the mean hourly weather for each calendar day across historical years, producing a normalized annual profile applied to all forecast years.
- **Rank-Based Averaging**: Sorts hourly weather within each month from coldest to hottest across historical years, averages by rank, and maps the ranked profile to a historical year's structure, preserving intra-month variability.
- **Multi-Year Weather Simulations**: Applies the forecast model to multiple historical weather years to generate a distribution of outcomes for each forecast year, enabling probabilistic interpretation (e.g., 50th percentile (P50) or 90th percentile (P90)). For the case study, P90 scenario was selected to align with 90/10 weather normalization practices commonly applied in the industry.

Each method is evaluated for its impact on forecast behavior, sensitivity to weather variability, and suitability for long-term distribution planning.

IV. RESULTS

This section presents the main outcomes of the case study analysis, focusing on forecast model performance and weather normalization impacts. These results provide insight into trade-offs between model design and normalization approach for time-series long-term forecasting in distribution planning. More detailed results, as well as additional findings not shown in this paper, can be accessed in the original report [15].

A. *Forecast Model Results*

Table 2 compares Models 1 and 5 using hourly MAPE, APE-Energy, and APE-Peak (annual, winter, and summer), reflecting common distribution planning priorities. MAPE captures hourly shape fidelity relevant for time-series analysis and DER studies, APE-Energy measures annual energy accuracy used in financial planning, and APE-Peak emphasizes

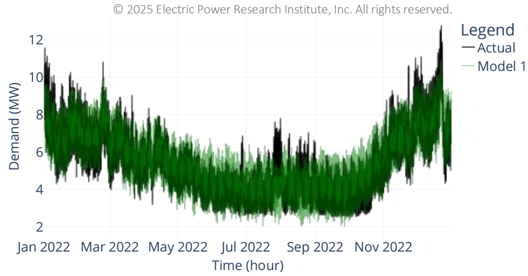
979-8-3315-8138-1

seasonal extremes that drive capacity upgrades and equipment sizing. Figure 4 also presents the hourly forecasted profiles for models 1 and 5, compared to the actual load for 2022, used as validation year. Model 1 tracks overall energy and shoulder-season levels, while Model 5 better captures winter peaks.

TABLE 2. PERFORMANCE METRICS FOR MODELS 1 AND 5

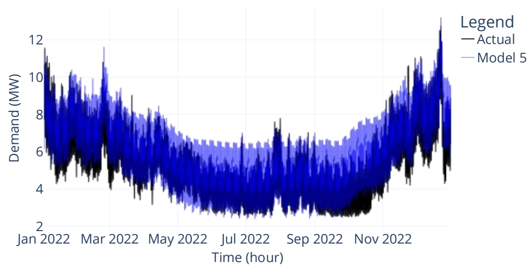
Forecast Model	Performance Metrics				
	MAPE [%]	*APE (Energy) [%]*	*APE (Peak) [%]*	*APE (Summer) [%]*	*APE (Winter) [%]*
Model 1	12.00	1.70	-15.40	-32.80	-15.40
Model 5	17.40	12.20	2.00	-7.40	2.00

© 2025 Electric Power Research Institute, Inc. All rights reserved.



© 2025 Electric Power Research Institute, Inc. All rights reserved.

(a) Model 1



(b) Model 5

© 2025 Electric Power Research Institute, Inc. All rights reserved.

Figure 4. Hourly forecasted profiles for models 1 and 5

Model 1, which uses continuous weather variables (TV), achieved the lowest MAPE and APE-Energy errors. These inputs provide smooth relationships between temperature and load, enabling accurate representation of average behavior and annual energy trends. However, Model 1 underperforms on peak metrics because it does not strongly weight extreme temperature conditions.

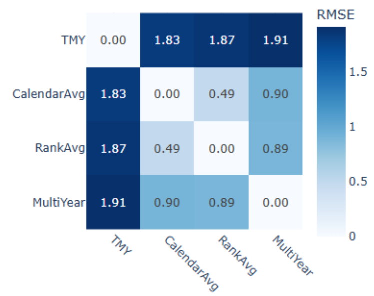
In contrast, Model 5 incorporates degree-hour variables (HDH24 and CDH24) and excels at predicting peak demand, particularly winter peaks. Degree-hour features emphasize extreme conditions but introduce many zero values outside heating and cooling periods, reducing accuracy for intermediate load levels. This trade-off improves peak prediction but sacrifices overall energy performance.

Overall, Model 1 is best suited for applications focused on energy, while Model 5 is preferable for capacity planning driven by peak conditions. Future work will explore hybrid approaches and advanced feature engineering to achieve strong performance on both energy and peak metrics within a unified modeling framework.

B. *Weather Normalization Results*

This subsection compares the four weather normalization methods applied to long-term feeder-level load forecasts for the period 2020–2035. The evaluation uses RMSE heatmap (Figure 5) to quantify the similarity among normalized profiles over the forecast horizon.

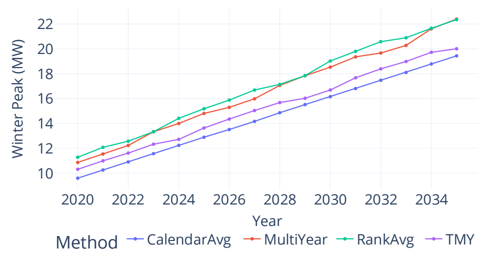
Results indicate that calendar-day averaging, rank-based averaging, and multi-year simulations form a consistent cluster, with RMSE differences below 1.0, suggesting these methods capture similar normalized weather/load behavior. Calendar averaging emerges as the most stable baseline, showing the smallest deviations compared to other approaches. In contrast, TMY stands out as an outlier, with RMSE values nearly twice as large relative to the other methods, indicating substantial divergence in forecast profiles. This discrepancy underscores the limitations of relying on a single composite year for feeder-level studies.



© 2025 Electric Power Research Institute, Inc. All rights reserved.

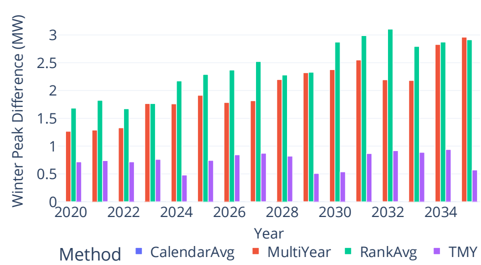
Figure 5. RMSE heatmap among the four methods

Annual energy and seasonal peak forecasts for the 2020–2035 horizon were also generated using the four weather normalization methods [15]. Figure 6 compares winter peak forecasts across these approaches, with line charts showing the trajectory of peak loads and bar charts illustrating differences relative to Calendar Averaging.



© 2025 Electric Power Research Institute, Inc. All rights reserved.

(a) Winter peak trajectory



© 2025 Electric Power Research Institute, Inc. All rights reserved.

(b) Winter peak difference

Figure 6. Winter peak forecast comparison among the four methods

979-8-3315-8138-1

Calendar Averaging consistently yields the lowest peak load forecasts, supporting year-over-year comparability, while Rank-Based Averaging produces the highest winter peaks by better preserving extreme cold conditions. These results highlight the trade-off between stability and risk representation: historical-based methods (Calendar, Rank) offer consistency, whereas probabilistic approaches (Multi-Year) provide a broader envelope for planning under uncertainty.

V. Discussion

The comparative analysis of normalization methods reveals that calendar-day and rank-based averaging provide stable, comparable results for year-over-year planning, while multi-year simulations offer a broader risk envelope for scenarios requiring probabilistic assessment. The choice of weather variable also plays a critical role in balancing energy and peak accuracy, underscoring the need to tailor model design to specific planning objectives. The analysis is based on a single representative feeder, and extension to feeders with different climates, customer mixes, or DER penetration is an important area for future work. In warmer climates or feeders with significant DER penetration, weather feature importance may shift, but the comparative behavior of calendar, rank-based, and multi-year normalization methods generally remains consistent, as their differences primarily reflect how weather variability is represented rather than feeder-specific characteristics.

Beyond weather normalization, broader research by EPRI addresses additional distribution planning forecasting challenges, including the alignment of system-level and distribution-level forecasting, and methodologies to allocate system-level forecast information to the feeder level. Such forecast allocation methodologies are essential for integrating system-level forecasts with local feeder projections, ensuring consistency and reliability across planning horizons. More detailed results and further discussion of these topics are available in the full report [15], supporting planners as they adapt to evolving grid conditions and increasingly complex forecasting requirements.

VI. Conclusion

Weather normalization is essential for producing reliable feeder-level load forecasts, particularly as distribution systems evolve under the influence of electrification and DER adoption. This paper explores how conventional peak-load focused weather normalization approaches can be extended to develop annual 8760-hour forecasts, which are increasingly being applied in advanced distribution planning studies. The findings highlight the importance of selecting normalization methods and weather variable designs that align with the intended planning use case, whether focused on energy, peak capacity, or risk assessment. Regular validation against actual data and ongoing refinement of model inputs is recommended to maintain forecast accuracy.

There are several promising future research directions, including automated diagnostics for feeder weather sensitivity, guidance on optimal history length for training data, and the development of advanced modeling techniques such as machine learning, hybrid models, and advanced feature engineering. Additional future research directions include addressing the integration of forecast allocation and alignment strategies and the evaluation of forecasting requirements for diverse distribution planning applications. As the grid continues to evolve, these advancements will help ensure that forecasting practices remain robust, adaptable, and aligned with emerging industry needs.

References

[1] T. Hong, J. Wilson and J. Xie, "Long Term Probabilistic Load Forecasting and Normalization With Hourly Information," *IEEE Transactions on Smart Grid,* vol. 5, no. 1, pp. 456-462, Jan. 2014.

[2] PJM, "Manual 19: Load Forecasting and Analysis. Sec. 4: Weather Normalization.," 18 Dec. 2024. [Online]. Available: https://learn.pjm.com/-/media/DotCom/documents/manuals/m19.ashx.

[3] R. Yang and J. Homer, "Load forecasting with climate variability for transmission and distribution system planning," 1 Nov. 2021. [Online]. Available: https://eta-publications.lbl.gov/sites/default/files/combined_pnnl_and_nrel_load_and_der_forecasting_ncep_fin.pdf.

[4] EPRI, "Load Forecasting Initiative (LFI)," [Online]. Available: https://msites.epri.com/lfi.

[5] "Forecasting in Distribution Planning: Evolving Techniques for Mid-Term and Long-Term Forecasting," EPRI, Palo Alto, CA: 2023. 3002026997.

[6] "Load Forecasting for Planning Timescales: 2024 Survey," EPRI, Palo Alto, CA: 2024. 3002030642.

[7] "Utility Experiences and Trends Regarding Data Centers: 2024 Survey," EPRI, Palo Alto, CA: 2024. 3002030643.

[8] "Accounting for Building and Transportation Electrification in Distribution Planning Forecasts: Existing Industry Practices, Capabilities and Gaps," EPRI, Palo Alto, CA: 2024. 3002030194.

[9] "Load Forecasting for Planning Timescales: Aligning Transmission and Distribution Forecasts," EPRI, Palo Alto, CA: 2024. 3002031259.

[10] J. S. McMenamin, "Defining Normal Weather for Energy and Peak Normalization. Itron Forecasting," 2008. [Online]. Available: https://na.itron.com/o/commerce-media/accounts/-1/attachments/3827793.

[11] E. S. I. Group, "Long-Term Load and DER Forecasting," 2025. [Online]. Available: https://www.esig.energy/long-term-load-and-der-forecasting/.

[12] "Distribution Planning Guidebook for the Modern Grid: 2024 Edition," EPRI, Palo Alto, CA: 2023. 3002030130.

[13] E. Fox, M. Russo, F. Monforte and S. McMenamin, "2022 PJM Model Review Final Report," 2022. [Online]. Available: https://www.pjm.com/-/media/DotCom/planning/res-adeq/load-forecast/pjm-model-review-final-report-from-itron.pdf.

[14] "Load Masking in Planning Forecast: Assessment of PV Modeling and Load Disaggregation," EPRI, Palo Alto, CA: 2021. 3002021439.

[15] "Load Forecasting for Distribution Planning: Accounting for Electrification and Weather in Time-Series Load Forecasts," EPRI, Palo Alto, CA: 2025. 3002033140.

979-8-3315-8138-1