Leveraging Short-Term Solar Forecasts for Load Disaggregation and Forecasting to Improve Distribution Operator Visibility

Technical Brief — *Distribution Operations and Planning*

Abstract

Distribution operators are facing challenges from load masking, where limited visibility of Distributed Energy Resources (DERs) can obscure the true (aka native or gross) load on the distribution system. While short-term forecasting is a cost-effective and scalable solution to providing DER visibility, a key question is how best to leverage DER forecasts to counteract load masking, thereby ensuring safe, reliable, and cost-effective operations of the distribution grid. This document details key learnings from a study with a large US utility and commercial forecast providers, with the goal of combining advanced solar and load forecasting techniques to estimate the current – and forecast the future – gross load for use in Distribution operations. Practical methods for real-time estimation (via load disaggregation) and forecasting are presented and validated against historical measurements. The presented methods are shown to be viable and capable of application at granularities from feeder-level down to customer/nodal-level.

Keywords

Short-term forecasting
DER visibility
Load disaggregation
Distribution operations
Grid operations

Executive Summary

PRIMARY AUDIENCE: Distribution operators

SECONDARY AUDIENCE: Distribution planners, Transmission operators

KEY RESEARCH QUESTION

This document aims to provide guidance on leveraging short-term forecasts to counteract load masking, with a focus on practical solutions that can be incorporated into Distribution operations. It further addresses how to use SCADA and AMI measurements to provide operational visibility of disaggregated load (gross load) in both real-time and operational planning timeframes This also includes validation of the proposed methods on a comprehensive dataset that includes contains net load, photovoltaic (PV) production, and gross load actuals.

RESEARCH OVERVIEW

The analysis evaluated the use of advanced forecasting techniques to provide customer and feeder-level estimates of gross load for scenarios where the net load is known from measurements, but not the distributed PV production. Performance was evaluated against real-world sites with distributed PV, with an emphasis on timescales relevant to Distribution operations (minutes to days ahead).

KEY FINDINGS

- Short-term PV forecasts can be used together with net load measurements to estimate disaggregated load (gross load) for "real-time" operations, without requiring PV measurement data. This includes estimates down to the customer-level, though estimates are more accurate when aggregated to coarser granularities. For example, customer-level estimates achieved errors of 6.4% ± 2.6%, while feeder-level estimates were 4.8%, relative to the peak load.
- Short-term PV forecasts can also be used to forecast gross load over timescales of interest to Distribution operations, without requiring PV measurement data. As with real-time estimation, forecasting can be done down to the customer-level, though forecasts are more accurate when aggregated to coarser granularities, e.g., feeder-level, due to diversity benefits.
- Gross load forecasts can be created by either a) first forecasting net load and DER production separately, then disaggregating to get a gross load forecast ('indirect'), or b) first disaggregating to estimate past gross load and then training a model to predict gross load ('direct'). Generally, the direct method results in more accurate forecasts, but there may be practical reasons to prefer the indirect method in Distribution operations, such as simplicity of implementation and transparency of results.
- Improving the accuracy of input DER forecasts enables improvements in the accuracy of gross load estimation and forecasting, all else being equal.

Table of Contents

This technical brief was prepared by EPRI.

WHY THIS MATTERS

The continued growth of DERs provides challenges to ensuring safe, reliable, and cost-effective operation of the distribution grid. These challenges are exacerbated by limitations on the visibility of DER production, which can mask the true (aka native or gross) load. This document provides insights and guidance on leveraging short-term DER forecasts for load disaggregation and forecasting. The learnings here can potentially be aggregated to the transmission level or used in Distribution planning, although this is not the focus of the paper.

HOW TO APPLY RESULTS

Distribution operators may use this report as a starting point on the integration of short-term solar and load forecasting, particularly in applications where there is a lack of DER visibility. The results may also be used as benchmarks for expected performance in terms of estimation and forecast accuracy.

LEARNING AND ENGAGEMENT OPPORTUNITIES

- This work is being conducted as part of a New York State Energy Research and Development Authority project. Future reports from this project will describe the application of these efforts.
- EPRI continues to collaborate with members and other industry stakeholders to improve the development and integration of distributed energy resources forecasting and load disaggregation. For utilities in the New York State area, engaging with this effort would allow use of the information in this report in grid operations.

EPRI CONTACTS: Mobolaji Bello, Technical Leader, mbello@epri.com
David Larson, Engineer/Scientist III, dlarson@epri.com

PROGRAM: Distribution Operations and Planning

IMPLEMENTATION CATEGORY: Reference

Introduction

Short-term forecasting can be used to address challenges related to Distributed Energy Resource (DER) visibility for Distribution operations. Figure 1 illustrates how lack of DER visibility can lead to uncertainty regarding the gross load, aka the 'native' load. Without knowledge of DER production (i.e., no DER visibility), a Distribution operator cannot conclusively determine the 'true' gross load. This lack of visibility can be particularly challenging with the growth of behind-the-meter (BTM) and small-scale DER systems (e.g., <50 kW), which are rarely equipped with monitoring equipment due to a variety of factors. But by using estimates, e.g., forecasts, of DER production in lieu of direct measurements, an operator can narrow down the range of possibilities and estimate the gross load. The question then is how to integrate short-term DER and load forecasting for use in Distribution operations.

This document aims to address the challenge of integrating forecasts of DER production vs load to maintain safe and reliable operation of the distribution grid, with learnings based on analyses using real-world DER installations in the US. For more details on the challenges of DER visibility and methods for short-term DER forecasting, readers are referred to the companion Technical Brief "Short-Term Forecasting for DER Visibility in Distribution Operations" [1].

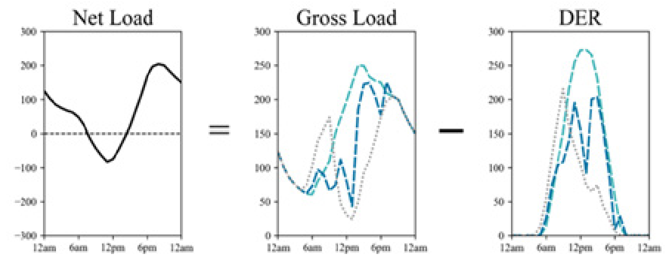

Figure 1. Example of how a single net load profile can be the result of any number of gross load and DER profile combinations.

The rest of the document is organized as follows. Utility Field Experience overviews sites and datasets used as the basis for this report. Load Disaggregation covers how to leverage short-term DER forecasts to disaggregate net load and estimate gross load, including benchmarks of estimation accuracy at the customer and feeder-level. Gross Load Forecasting builds upon the load disaggregation approach to then predict future gross load for use in Distribution operations. Finally, Conclusions summarizes the key takeaways from the work and future research directions.

Utility Field Experience

This document is based on learnings from on-going work between EPRI, a large US utility (hereafter referred to as Utility), a commercial solar forecast provider (hereafter referred to as Solar Forecaster), and a commercial load forecast provider (hereafter referred to as Load Forecaster). The details provided here will be referenced throughout the rest of this document. Note that the work presented here is from the same project referenced in [1].

Sites

The Utility, in coordination with EPRI, identified over 70 sites for use in the study. The selection of sites was based on the following criteria: (1) sites with existing small-scale (<50 kW) PV and BTM installations, (2) sites with co-located meters for (separately) measuring net load and PV production, and (3) sites located within the same region of the Utility's service territory. In total, the selected sites represent over 400 kW of PV nameplate capacity.

Datasets

For each site, the Utility provided approximately two years (2019–2020) of hourly-resolution net load and 15-min resolution PV production actuals. EPRI aggregated the site-level net load and PV production data to create 'pseudo' feeder-level datasets. In parallel, the Solar Forecaster provided 15-min resolution PV production forecasts for each site covering horizons of minutes to days ahead. Additionally, both the PV and load forecasts were created in accordance with operational constraints, i.e., the models only used information that would be available at the time of fore-

Table 1. Forecast horizon vs use cases for Distribution operations.

Horizon	FLISR	Hosting Capacity	VVO	DER Dispatch	Maintenance/ Scheduling	Contingency Analysis
Nowcast	✓	✓	✓			✓
Intra-hour	✓	✓	✓	✓		✓
Intra-day		✓	✓	✓		✓
Day-ahead		✓	✓	✓		✓
Week-ahead					✓	✓
Month-ahead					✓	✓

cast. Note that the PV production forecasts were generated using a methodology that does not rely on measurements of PV production and instead only requires standard information regarding the PV system details, e.g., installed capacity, DC:AC ratio, panel orientation and tilt. See [1] for more details on the PV forecast methodology.

Note: unless stated otherwise, all reported results are based on the hourly-resolution data from Summer 2020 (June–September). The data from January 2019–May 2020 was used for training the load forecast models and therefore was not included when evaluating forecast model performance.

Forecast Use Cases

Table 1 summarizes forecast horizon vs use cases for short-term forecasts in Distribution operations that were identified by EPRI in collaboration with the Utility:

- Fault Location, Isolation and Service Restoration (FLISR)
- Dynamic Operating Envelopes for Constraint Management
- Volt-Var Optimization (VVO) with DER (smart inverters)
- DER dispatch, e.g., energy storage (ES)
- Maintenance/scheduling
- Contingency analysis

where the horizons are classified here as nowcasts (<5-mins ahead), intra-hour (<1-hour), intra-day (≥1 hour), day-ahead (≥24 hours), week-ahead (≥7 days) and month ahead (≥30 days). Note however that this is not an exhaustive list and that integrating forecasts in Distribution operations is an area of active research.

Load Disaggregation

The first task was to focus on the problem of estimating gross load–aka the native load–when the net load is known but not the DER production. This problem can be thought of as a specialized case of non-intrusive load monitoring (NILM), where the goal is to disaggregate a (net) load signal into its underlying components. For example, given AMI measurements of load of a home, NILM methods could be used to estimate what percentage of the aggregate load is due to a refrigerator vs washing machine vs air conditioning. However, in the task evaluated, the concern is solely with disaggregating DER production from (measured) net load to get the gross load, which we will refer to as 'load disaggregation.'

One solution to estimating gross load is to use short-term forecasts as a proxy for (direct) measurement of DER production. Assuming net load measurements are available and that the forecasts are sufficiently accurate, the gross load can then be estimated as:

$$\text{Gross Load} = \text{Net Load} + \text{DER Production},$$

which is a simple rearrangement of:

$$\text{Net Load} = \text{Gross Load} - \text{DER Production}.$$

Figure 2 provides a visual explanation of this approach to load disaggregation.

The challenge then is in estimating the DER production without access to measurements of DER production. In the case where the DER is PV without energy storage, the PV production can be estimated using physical modeling based methods and details regarding the PV system, e.g., nameplate capacity, DC:AC ratio, panel orientation and tilt. For example, solar irradiance at any location in the US can be estimated from geostationary satellite imagery, which can then be translated into PV production using commercial or open-source irradiance-to-power modeling tools, e.g., PVLIB [2] or the System Advisor Model (SAM) [3]. For other DER, further assumptions must be made regarding the DER operation and therefore purely physical modeling based methods may not be viable for estimated DER production. For example, a battery energy storage system (BESS) could be operated for demand-charge reduction, energy arbitrage, demand response or a number of other applications, each resulting in a different DER production profile despite being based on the same DER technology.

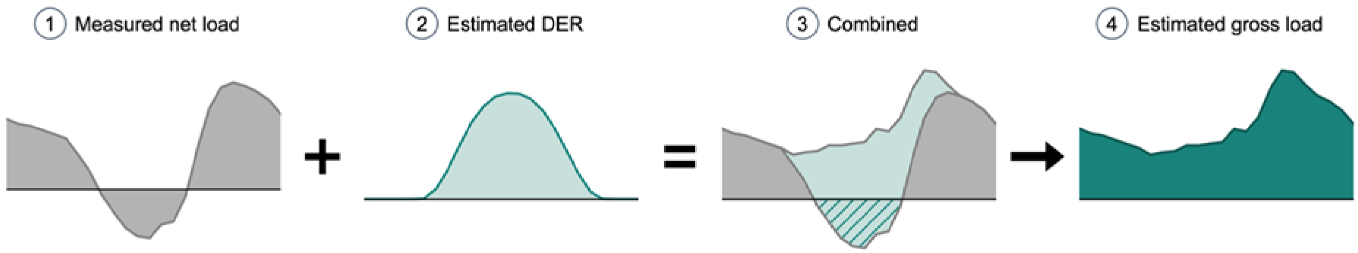

Figure 2. Illustration of the gross load estimation method.

Estimation Accuracy

The proposed load disaggregation approach for estimating gross load can be used regardless of the spatial granularity. This commenced with site-level estimation using hourly-resolution net load measurements from AMI and PV production nowcasts (<5-min ahead) provided by the Solar Forecaster, with the gross load actuals computed from net load and PV production measurements provided by the Utility. Figure 3 shows the accuracy of the gross load estimation–in terms of root mean square error (RMSE) normalized by peak load–as a function of the PV nameplate capacity and peak load per site. Excluding one outlier site (with a RMSE of 62.5%)[1], the RMSE values are fairly uniform (6.4% ± 2.6%), indicating that estimation accuracy is not (strongly) dependent on peak load, PV capacity or other variations between sites.

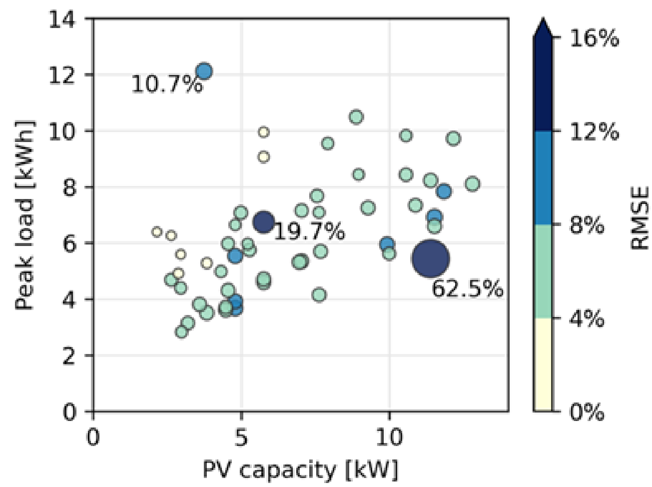

Figure 3. Site-level gross load estimation error (marker color) as a function of PV capacity and peak load. Each circle marker represents one site.

Applying the same approach to estimate the feeder-level gross load results in a RMSE of 4.8%. The lower error (higher accuracy) of the feeder-level estimation can be primarily attributed to two complementary factors. First, aggregating PV production from spatially distributed sites tends to reduce variability due to spatial smoothing effects and lower variability PV production is easier to predict. Second, aggregating forecasts to coarser granularities can be more accurate due to errors from finer granularity forecasts canceling each other out such that the overall error is reduced.

Figure 4 compares the daily profiles of gross load calculated from measurements vs estimated using forecasts. While the profiles are not identical, the figure illustrates how the proposed approach can also be used to develop load profiles for use in planning studies, though that is beyond the scope of this document.

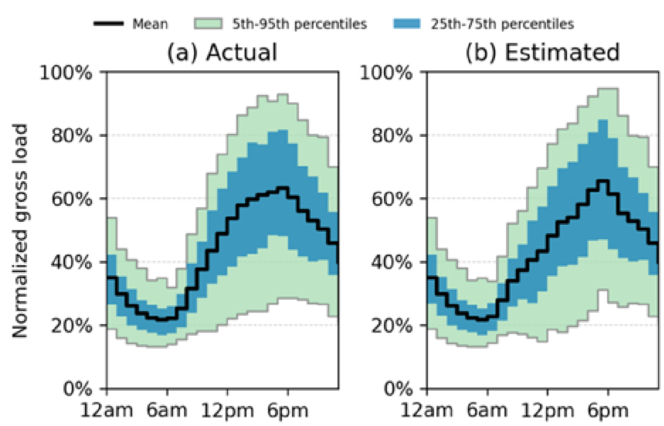

Figure 4. Distribution of feeder-level gross load by hour of the day using net load measurements with either (a) measurements of PV production or (b) estimated PV production from forecasts.

Gross Load Forecasting

Next is addressing the forecasting problem. More specifically, how to leverage short-term PV forecast when forecasting gross load. Here the goal is to make a prediction at time T of the gross load at future times $\{T+1, T+2, \ldots\}$ using only information available at time T or before $\{T-1, T-2, \ldots\}$. Two approaches are considered, which are referred to as 'indirect' and 'direct' (as illustrated in Figure 5). Unless stated otherwise, all

[1] The outlier site had uncharacteristically large PV forecast errors, which were due to incorrect PV system details (e.g., AC capacity). The site was left in the analysis to highlight how results depend on the accuracy of the PV system details and that measurements can be used to diagnose model performance.

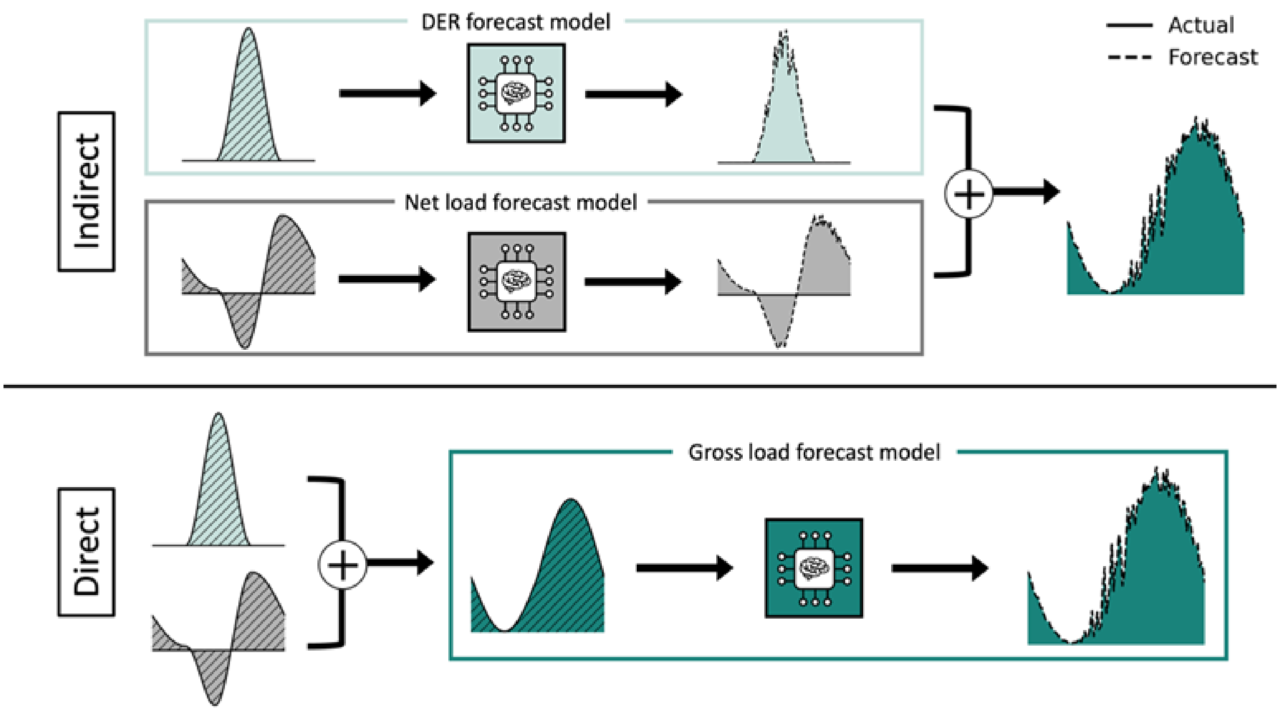

Figure 5. Comparison of indirect vs direct forecasting approaches. The indirect approach involves first forecasting net load and DER production (separately), and then disaggregating to get the gross load forecasts. In contrast, the direct approach first disaggregates the past load and then uses a statistical model to predict the gross load.

results are based on a schedule where hourly-resolution forecasts are generated once per day at midnight (local time) for the next 24 hours.

The indirect method involves first forecasting the net load for future times {*T*+1, *T*+2, …}, and then disaggregating the net load forecast using a DER forecast to predict the gross load for {*T*+1, *T*+2, …}. In contrast, the direct method starts by estimating the gross load for the prior times {*T*-1, *T*-2, …} by disaggregating the past net load and then applying a model to directly predict the gross load for future times {*T*-1, *T*-2, …}. In other words, either (a) forecast net load first and then disaggregate to predict the gross load (indirect method) or (b) disaggregate the past net load first and then forecast using the estimated past gross load as input (direct).

Forecast Accuracy

Figure 6 shows the accuracy of the indirect method for site-level forecasting using PV production forecasts from the Solar Forecaster and net load forecasts from the Load Forecaster using Machine Learning models trained on historical net load data. Similar to the gross load estimation results (Figure 3), the forecast accuracy (normalized RMSE) is fairly uniform across sites, aside from the same outlier site. And at the feeder-level, the indirect method achieves a RMSE of 12.2%. While the forecast error magnitudes are greater than the gross load estimates, this is an expected result given that forecast accuracy tends to decrease as the horizon increases.

The direct method of forecasting gross load–using PV production forecasts from the Solar Forecaster and net load measurements from the Utility–shows similar trends to the indirect method, i.e., fairly uniform accuracy across sites and higher RMSE values than the gross load estimation results. However, the direct method results in lower error for both site-

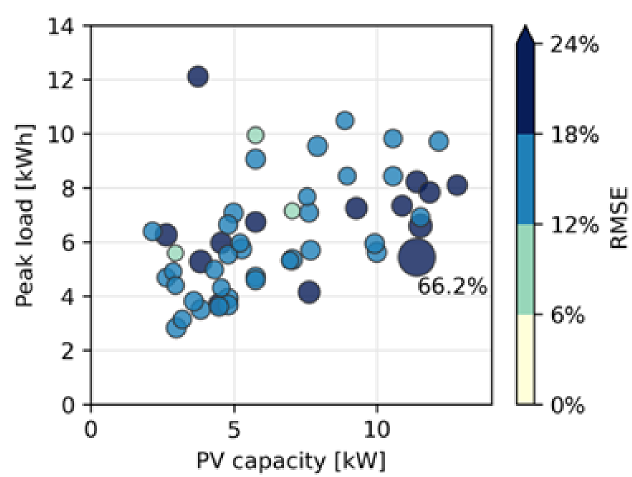

Figure 6. Site-level gross load forecast RMSE (marker color) as a function of PV capacity and peak load. Each circle marker represents one site.

level (15.9% ± 7.2% vs 17.3% ± 7.5% for the indirect method) and feeder-level forecasts (9.8% vs 12.2% for the indirect method). These results have parallels to transmission level load forecasting, including the use of short-term forecasts of distributed PV to counteract lack of DER visibility [4].

The higher accuracy (lower error) of the direct method can be attributed to the two factors. First, training of the Machine Learning model allows the direct method to minimize error by tuning the model using historical data, whereas the indirect method (as implemented) assumes the input net load and PV production forecasts are already 'optimal'. Therefore, the indirect method can lead to compounding errors, e.g., over-prediction of net load at the same time as over-prediction PV production would result in even greater over-prediction error. Second, gross load tends to be less variable than net load with PV production and therefore easier to predict. However, note that while these two factors are likely to apply to other systems, leading to the direct method (generally) resulting in lower error than the indirect method, the results may vary significantly between distribution grids due to differences in, e.g., weather/climate, penetration of DER and customer behaviors.

Impact of DER Forecast Accuracy

While the direct method achieves lower error than the indirect method using the same PV production forecasts, a natural follow-up question is whether the error can be further decreased by improving the PV production forecasts. Direct forecasts produced are compared using the same procedure, with the only change being the input PV production forecasts:

- **Baseline:** same PV production forecasts as the prior subsection; PV forecasts generated once per day at midnight (local time) for the next 24-hours hours
- **Improved:** PV production nowcasts (<5-min ahead), which represent an upper bound on the accuracy of commercially available PV production forecasts
- **Perfect:** measurements of PV production, which represent PV forecasts with zero error and therefore a 'best case' scenario for gross load forecasting

Figure 7 compares the distribution of site-level RMSE for the indirect method and three variations of the direct method. Direct forecasting using the improved PV production forecasts reduces the mean RMSE but has a negligible impact on the spread of RMSE values (14.9% ± 6.9% vs 15.9% ± 7.2% for direct baseline). In comparison, direct forecasting with perfect PV production forecasts further reduces the mean RMSE while also significantly reducing the RMSE spread (12.9% ± 3.2%). A similar trend is observed for feeder-level forecasts, with the choice of PV production forecasts reducing the RMSE from 9.8% (baseline) to 7.3% (improved) to 6.3% (perfect). The reduction in error indicates that the direct method accuracy is dependent on the PV production forecast accuracy, but only to a limit.

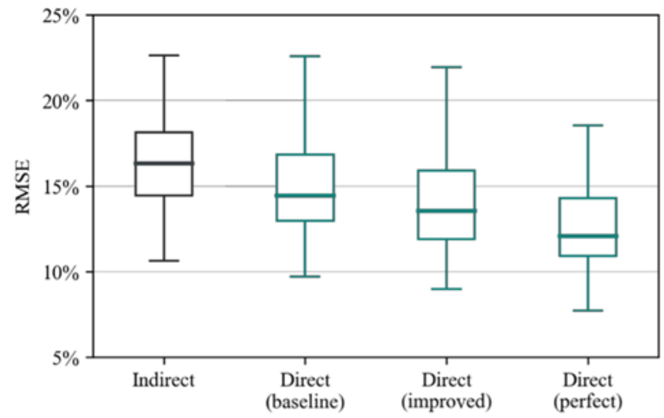

Figure 7. Site-level gross load forecast RMSE for the indirect method vs the direct method with three levels of input PV forecast accuracy (baseline, improved and perfect).

Considerations Beyond Forecast Accuracy

The accuracy results imply that the direct method is preferable over the indirect method when the primary goal is minimizing gross load forecast error. However, there are practical considerations that may influence a Distribution operator to adopt the indirect method despite its (comparatively) lower accuracy. Here, some of these factors are briefly discussed, while emphasizing that the importance of these factors and their influence on the final decision may vary significantly between Distribution operators, as well as change over time.

- **Simplicity of implementation:** The indirect method requires overlapping forecasts of both net load and DER production but is otherwise agnostic to how the input forecasts are produced and only requires adding the two forecasts, which makes the method implementable in almost any scripting language (Python, Matlab, etc.), Excel spreadsheets or even directly via a database query. In contrast, the direct method requires coordination between the DER and load forecasters, who can be different entities.
- **Transparency of results:** The simplicity of the indirect method (i.e., gross load forecast = net load forecast + DER production forecast) also means it may be viewed as clearer to forecast users, particularly in terms of the causes of forecast errors. In contrast, the direct method may be viewed as more of a "black box" and therefore less directly informative to Distribution operators.
- **Robustness and scalability:** The separation of the net load and DER production forecasts means the indirect method can immediately switch to improved or alternative forecasts as they become available, without any retraining. This also means the indirect method may be able to adapt more quickly to changes in a system, e.g., additional DER capacity coming online, and use multiple forecasts in an ensemble, which is a common practice in transmission level forecasting.

Conclusions

Distribution operators are facing challenges related to DER visibility and load masking as the growth of PV and other DERs continue to increase. A potential solution to these challenges–that is both cost-effective and scalable–is to leverage short-term forecasting techniques. Results from on-going work with a larger US utility and commercial forecast providers have resulted in several key takeaways on leveraging short-term forecasting for Distribution operations:

- Nowcasts (<5-min ahead) of PV production can be combined with net load measurements to provide "real-time" estimation of gross load in scenarios where the PV production (and therefore gross load) is unknown. At the feeder-level granularity, gross load estimation error can be as low as 4.8% of the peak load, compared to 6.4% ± 2.6% for customer-level estimates.
- Gross load forecasts can be created by either a) forecasting net load first and using a DER forecast to disaggregate to get the gross load (indirect) or b) disaggregate past net load to estimate past gross load and then forecasting using the past gross load as input (direct). Results indicate the direct forecast approach achieves higher accuracy (lower error) overall.
- The accuracy of both gross load estimation and forecasting can be increased by improving the input DER forecast data, all else being equal.
- Gross load estimation and forecasting at coarser granularities, can be more accurate, e.g., aggregating site-level estimates to produce feeder-level estimates can yield higher accuracy (lower relative error) than individual site-level estimates. But site-level estimates and forecasts may still be sufficiently accurate for use in Distribution operations.
- While not the focus here, similar methods could be leveraged to provide forecasts of aggregate DER to bulk system use cases, realizing that the granularity needed would likely be different.

The next step is to build upon the learnings from this document to focus on integrating short-term forecasts into Distribution operations. A follow-up technical brief is planned that will demonstrate use cases for short-term forecasts in Distribution operations, e.g., FLISR, dynamic operating envelopes for constraint management, VVO, and DER dispatch. This is expected to include benchmarks and guidance for Distribution operators looking to incorporate short-term forecasting.

References

[1] "Short-Term Forecasting for DER Visibility in Distribution Operations" (3002021931) available at https://www.epri.com/research/products/000000003002021931.

[2] Holmgren, W. F., Hansen, C. W., and Mikofski, M. A., "PVLIB Python: A Python Package for Modeling Solar Energy Systems," Journal of Open Source Software, 3(29), 884, 2018.

[3] Blair, N., DiOrio, N., Freeman, J., Gilman, P., Janzou, S., Neises, T., and Wagner, M. "System Advisor Model (SAM) General Description (Version 2017.9.5)," NREL/TP-6A20-70414.

[4] Black, J. "Short-term load forecasting at ISO/RTOs with high penetration of distributed PV", ESIG Blog, February 23, 2017. https://www.esig.energy/blog-short-term-load-forecasting-isortos-high-penetration-distributed-pv/.

Contact Information

For more information contact the EPRI Customer Assistance Center at 800.313.3774 (askepri@epri.com).

EPRI Resources

Mobolaji Bello, *Technical Leader*
865.218.8005, mbello@epri.com

David Larson, *Engineer/Scientist III*
650.855.2318, dlarson@epri.com

The authors would like to thank Lysha Matsunobu, Aidan Tuohy, and Lindsey Rogers for their contributions to this document.

Distribution Operations and Planning

The Electric Power Research Institute, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI also provides technology, policy and economic analyses to drive long-range research and development planning, and supports research in emerging technologies. EPRI members represent 90% of the electricity generated and delivered in the United States with international participation extending to nearly 40 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; Dallas, Texas; Lenox, Mass.; and Washington, D.C.

Together . . . Shaping the Future of Energy™

3002022915 November 2021

EPRI
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com