

UC Berkeley

UC Berkeley Previously Published Works

Title

Getting brighter: Impacts of improved day-ahead solar forecasts in high-solar, high-storage electricity systems

Permalink

<https://escholarship.org/uc/item/3qw104d7>

Journal

Renewable Energy, 259

ISSN

0960-1481

Authors

Kahrl, Fredrich
Jorgenson, Jennie
Kiboma, Lawryn
[et al.](#)

Publication Date

2026-03-01

DOI

10.1016/j.renene.2025.125104

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

Getting Brighter: Impacts of Improved Day-Ahead Solar Forecasts in High-Solar, High-Storage Electricity Systems

Fredrich Kahrl,^{a,b,c} Jennie Jorgenson,^d Lawryn Kiboma,^d Dev Millstein,^a Natalie Mims Frick,^a Brian Sergi^d

^a Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720

^b University of California, Berkeley, Berkeley, CA 94720

^c 3rdRail Inc., 1225 Allston Way, Berkeley, CA 94720

^d National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401

March 2026

This is a preprint version of an article published in *Renewable Energy*.
DOI: <https://doi.org/10.1016/j.renene.2025.125104>



DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

COPYRIGHT NOTICE

This manuscript has been authored by an author at Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231 with the U.S. Department of Energy. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges, that the U.S. Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

Getting Brighter: Impacts of Improved Day-Ahead Solar Forecasts in High-Solar, High-Storage Electricity Systems

Fredrich Kahl,^{a,b,c,*} Jennie Jorgenson,^d Lawryn Kiboma,^d Dev Millstein,^a Natalie Mims Frick,^a Brian Sergi^d

^a Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720

^b University of California, Berkeley, Berkeley, CA 94720

^c 3rdRail Inc., 1225 Allston Way, Berkeley, CA 94720

^d National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401

* Corresponding author, fkahl@berkeley.edu

Abstract

This paper analyzes the impacts of improved day-ahead solar forecasts on costs and dispatch in the solar-rich Southeast U.S. It uses a high-solar, high-storage resource portfolio in which solar generation capacity accounts for 45% of total installed capacity (34%-36% of generation) and energy storage capacity (43 GW) is equivalent to 33% of peak demand. In a base scenario, improved day-ahead solar forecasts reduce production costs by \$87 million per year (\$0.13 per MWh load, 2023\$). This level of savings is within the range or lower than earlier studies of solar forecast improvements at lower levels of solar generation (< 25% of total generation). In this study, solar expansion was accompanied by two important sources of flexibility for managing solar forecast error: energy storage and day-ahead solar curtailment. The analysis finds that regional coordination complements day-ahead solar forecast improvements while natural gas commitment flexibility is a substitute for forecast improvements. Day-ahead solar forecast improvements reduce reserves required to manage forecast error by 30%. Fewer reserves to manage large, infrequent solar forecast errors could be an important benefit of improved solar forecasts.

Keywords: Solar PV, solar forecast, energy storage, system flexibility

1. Background

Since the mid-2010s a growing number of studies have examined the benefits of improved solar forecasts, from both solar plant operator and system operator perspectives. Studies from a solar plant operator perspective examined reduced penalties, imbalance costs, or reserves requirements from improved forecasts, covering diverse geographic contexts: Australia [1], California [2,3], Spain [4,5], and in U.S. organized markets [6]. These studies found significant benefits (lower penalties, imbalance costs, required reserves) to market participants that, in many cases, increased linearly with increases in solar forecast accuracy.

Studies from a system operator perspective have examined the change in system costs or market costs. Martínez-Anido et al. [7,8] examined the impacts of uniform (25% to 100%) improvements in solar forecasts for solar penetrations ranging from 4.5% to 18% in ISO-NE. The study found that improvements in day-ahead solar forecasts reduced real-time generation from faster start combustion turbines and steam plants, startup and shutdown costs, ramping, and curtailment, leading to production cost savings that ranged from \$0.11-\$1.42/MWh, with the range driven mainly by the level of forecast improvement and the level of solar generation. Total benefits increased with solar forecast improvements and solar generation levels, but there were diminishing returns to forecast improvements.

Hodge et al. [9] studied the impacts of solar and wind forecasting improvements (uniform 20% and 40% improvements) with different levels of wind and solar generation (0% to 56%, up to approximately 25% solar) and different levels of energy storage (100 MW/400 MWh and 200 MW/800 MWh) for a generic 118 bus system with San Diego Gas & Electric (SDG&E) loads and generation mixes from the California Independent System Operator (CAISO, higher gas) and Midcontinent Independent System Operator (MISO, higher coal) regions. The study found that improved solar and wind forecasts led to \$0 to \$7 million in annual production cost savings across scenarios (\$0.04-\$0.28/MWh, assuming an annual load of 25 TWh for SDG&E), with higher savings in higher renewable scenarios and in many cases with higher

forecast improvements. In most cases, storage reduced the value of renewable forecast improvements but only between the “no storage” and 100 MW scenarios.

Goodarzi et al. [10] analyzed historical data to examine the impact of wind and solar forecast errors on energy imbalances and spot market prices in Germany. The analysis found that higher wind and solar forecast errors tend to increase imbalances and spot prices, though wind forecast errors had a larger impact than solar forecast errors. Wang et al. [6] used 2012-2019 day-ahead and real-time market prices to calculate the costs of solar forecast errors in different U.S. independent system operator (ISO) markets, with day-ahead solar forecasts calculated using the North American Mesoscale (NAM) model and a persistence algorithm. The analysis found that the cost of solar forecast errors in different ISOs was generally less than \$1/MWh, that costs were higher in ISOs with more solar, and that solar forecast improvements had diminishing returns.

Hobbs et al. [11] used synchronous (concurrent with each hour in the same year) and asynchronous (errors taken from different times) solar reforecasts and a multi-stage production simulation model to examine the cost of solar forecast errors and the impact of forecast improvements in the Southern Company service territory with approximately 15% solar generation. The study found \$10 million in annual costs (\$0.06/MWh) from solar forecast errors, \$7 million in savings (\$0.04/MWh) from a 50% reduction in forecast cost error, and diminished returns to forecast error improvements. Cui et al. [12] studied the reliability benefits of improved solar forecasting on multiple time scales (day-ahead, hour-ahead, five-minute, sub-five-minute), focusing on reliability metric improvements rather than costs. The study found that a 10% improvement in solar forecast accuracy led to a 9%-13% reduction in control performance score (CPS) violations.

There has been little to no documentation of the cost of solar forecast improvements. Since the early 2010s there has been an understanding within the U.S. electricity industry that the cost of developing and implementing wind and solar forecasting systems would be small relative to benefits [13]. Drawing on U.S. examples, Gandhi et al. [14] report that thus far the cost of wind and solar forecasting systems has been low. For instance, CAISO charges wind and solar plant owners \$0.10/MWh (per MWh of wind and solar generation) to use its wind and solar forecasting systems [15], which implied forecasting costs per unit load of \$0.03/MWh in 2023 [16]. However, the cost of improving solar forecasts is less well-understood.

To summarize these studies, improvements in solar forecasts generally benefit both plant operators and the broader transmission system (system operator perspective), though as Luoma et al. [2] noted plant operators (market participants) may not always have an incentive to submit accurate forecasts in their market offers. From a system operator perspective, the cost of solar forecast error, and the benefit from improving solar forecasts, was between \$0.04/MWh-\$1.42/MWh with the resource mixes studied. The benefits of improved solar forecasts increased with higher levels of solar, but in most cases the systems studied had solar generation levels of less than 20%. More system flexibility appeared to reduce the benefits of improved solar forecasts in some studies [6,9], but the relationship between solar forecasts and flexibility was not studied systematically. Solar forecasts appeared to have a significant impact on storage operations in the Hodge et al. [9] study but the relationship between solar forecasting and storage operations was also not studied in depth.

This study focuses on a system operator perspective. It builds on prior studies but extends the analysis in several ways. First, it examines the operational impacts of an improved solar forecast in a system with

higher levels of solar (45% of installed capacity) and energy storage (43 GW, 33% of peak demand) than have been previously studied. It focuses on a solar-rich region (Southeast U.S.) that has very little wind generation but instead has a diverse mix of coal, hydro, natural gas, and nuclear generation. Second, this study considers base and improved solar forecasts that are structurally different from one other, rather than assuming uniform improvements in forecast accuracy. Third, it more explicitly studies the relationship between the benefits of solar forecast improvements and system flexibility, including the flexibility to import from and export to neighboring systems and the flexibility of natural gas generation. Fourth, it uses a more realistic characterization of day-ahead and real-time operations and reserves than in most previous studies, with hourly day-ahead scheduling that constrains 15-minute real-time dispatch and solar forecasts rather than solar profiles in day-ahead lookaheads.

2. Methods

This section describes the geographic focus of the study, the resource portfolios used in the analysis, methods for solar forecasts and forecast errors, methods for calculating forecast error reserves and reserve amounts, methods for production simulation, and the modeling scenarios used in the analysis.

2.1 Geographic Focus and Resource Portfolios

This study focuses on four regions in the solar-rich Southeastern U.S., covering Alabama, Georgia, Kentucky, North Carolina, South Carolina, and Tennessee. The four regions include DUK (Duke Energy Carolinas, Duke Energy Progress Energy Carolinas, Dominion Energy South Carolina, and South Carolina Public Service Authority), LGEE (Louisville Gas & Electric and Kentucky Utilities), SOCO (Southern Company), and TVA (TVA) (Figure 1). They approximately capture but do not match up perfectly with balancing (control) areas and thus we refer to them as ‘balancing regions’ in this analysis.

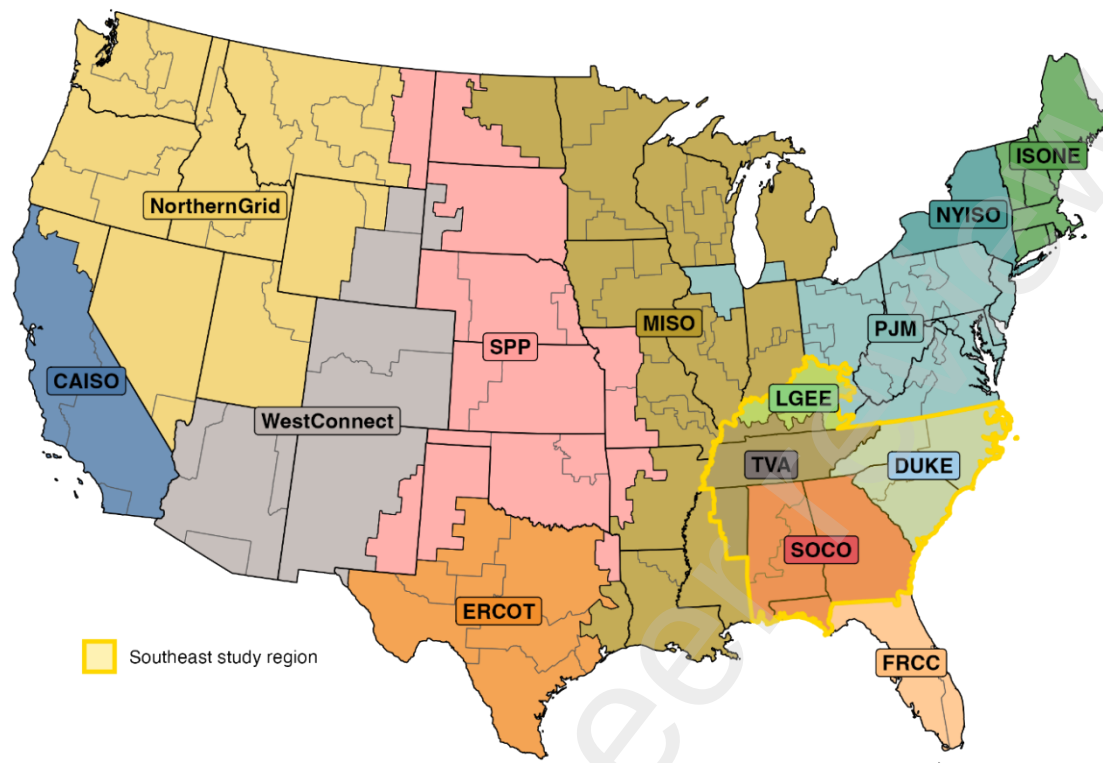


Figure 1. Map of Southeast Study Region

The resource portfolios we use in the analysis are from the medium PV high storage (MPHS) scenario in Kahrl et al. [17], which is a future year (2035) scenario. We chose the MPHS scenario because it has high levels of both solar PV and storage capacity, relative to current capacity, but still has a significant amount of coal, natural gas, and nuclear generation. The MPHS portfolios were developed for using the ReEDS model [18], a national capacity expansion model, and contain least-cost resource mixes that were developed using a consistent set of assumptions across all four regions. In total, the four regions have 137 GW of installed solar photovoltaic (PV) capacity (45% of total installed capacity) and 43 GW / 292 GWh of energy storage (33% of coincident peak demand, 0.04% of annual energy consumption). Appendix A.1 has additional detail on the resource portfolios.

2.2 Solar Forecasts

The analysis compares two day-ahead solar forecasts: (1) a less accurate forecast, using a simple day-ahead persistence (PST) forecast, and (2) a more accurate forecast, based on the North American Mesoscale (NAM) forecast system. NAM is a numerical weather prediction model produced by the U.S. National Centers for Environmental Prediction (NCEP) that is commonly used in solar forecasting studies [3,6,12]. We chose these forecasts to meet two main study design criteria: (1) use of publicly available data, (2) structural differences between the less and more accurate forecasts rather than uniform improvements. Persistence forecasts have been used as the base (less accurate) forecast in studies where the more accurate solar forecast is not based on uniform improvements [6].

We use NAM to develop forecasts and real-time solar profile data (see Appendix A.2). We calculate the simple day-ahead persistence forecast by assuming that available solar in hour h of a future day ($S_{d+n,h}$, n days into the future) is equal to the current day's available solar in hour h ($S_{d,h}$). For instance, the persistence forecast for a day-ahead forecast will be $S_{d+1,h} = S_{d,h}$, the second day in a 24-hour lookahead will be $S_{d+2,h} = S_{d,h}$, and so on (see Section 2.4 for a description of lookahead periods). For the NAM forecasts, we develop forecasts for day-ahead (24-hour-ahead) and the day-ahead lookahead periods (48 and 72 hours ahead of the forecast interval) using the methods described in Appendix A.2.

Although forecast accuracy can be difficult to compare across studies, forecast accuracy improvements in previous studies provide a useful benchmark for this study (Table 1). The forecast accuracy improvements in this study are on par with those in Hodge et al. [9] and Hobbs et al. [11] and are consistent with improvements in reported monthly solar forecast errors in CAISO, which were 3.5%-6.0% in 2018-2019 and improved to 2.2%-4.1% in 2021-2022 [19,20]. Zhang et al. [21] present a more extensive list of metrics for comparing solar forecasts and forecast improvements. Appendix A.3 provides additional forecast error results.

Table 1. Total Normalized Mean Absolute Errors (NMAE) and Normalized Root Mean Square Errors (nRMSE) in this Study Relative to Other Studies

	Metric	Value	
		Including zero values	Excluding zero values
This study			
Persistence forecast	NMAE	5.5%	9.7%
	nRMSE	10.8%	14.4%
NAM forecast	NMAE	2.8%	4.9%
	nRMSE	5.8%	7.7%
Hodge et al. [9]			
Base forecast	NMAE	2.0%-2.4%	
20% forecast improvement	NMAE	1.5%-1.9%	
40% forecast improvement	NMAE	1.1%-1.4%	
Wang et al. [6]			
Persistence forecast	nRMSE		15.0%-23.6%
NAM forecast	nRMSE		13.4%-20.8%
Hobbs et al. [11]			
Base forecast	NMAE		5.3%
Forecast error reduced by half	NMAE		2.7%

Notes: Total refers to Southeast region wide. In this study, we normalize errors by installed capacity rather than maximum generation. Including nighttime hours (zero solar values) reduces error metrics relative to excluding nighttime hours. NMAE metrics in Hodge et al. [9] include zero values. Wang et al. [6] exclude zero values. Hobbs et al. [11] did not specify whether zero values were included, though we assume here zero value periods were excluded as base NMAE (5.3%) roughly match the NMAE in the NAM forecasts in this study (4.9%).

2.3 Forecast Error Reserves

Increasing reserves to manage solar forecast error can improve system balance [22]. We hold three kinds of operating reserves in production simulation modeling, based on the approach in previous studies [23,24,25]: (1) day-ahead and real-time spinning and non-spinning contingency reserves, each 3% of load; (2) day-ahead spinning and non-spinning forecast error reserves to manage day-ahead solar

and wind forecast error, equivalent to 1 and 2 standard deviations of day-ahead forecast errors, respectively; and (3) day-ahead and real-time regulation reserves, equivalent to 1% of load plus reserves to manage 3 standard deviations of 15-minute forecast errors. For more on the reserve calculation methods, see [17].

We calculate forecast error-related reserves exogenously using solar and wind profile data. In the model, we hold all three kinds of reserves in the day-ahead simulations but release the day-ahead forecast error reserves (2) in the real-time simulations. We pool (share) day-ahead forecast error reserves (2) across the Southeast region due to the high level of reserve requirements, assuming they would be pooled in practice. We allow solar and wind generation to provide spinning and regulation reserves [24] and energy storage to provide all reserves.

Table 2 shows annual average day-ahead forecast error reserve (2) requirements. The average reserve levels in Table 2 are high relative to the amount of operating reserves that system operators currently tend to hold due to the high levels of solar generation in this study. Relative to the persistence forecast, the NAM forecast reduces spinning and non-spinning reserves required to manage day-ahead forecast error by around 30%.

Table 2. Annual Average Day-Ahead Load Following Reserve Requirements with Persistence (PST) and NAM Forecasts

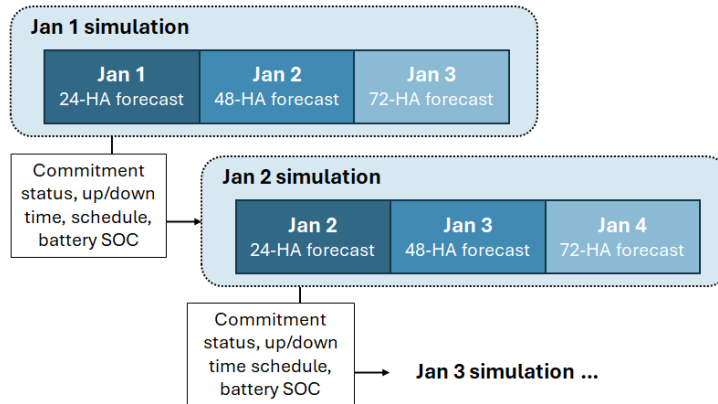
Forecast	Average spinning reserves (MW, % load)	Average non-spinning reserves (MW, % load)
PST	12,766 (16%)	25,533 (33%)
NAM	8,789 (11%)	17,579 (22%)

Notes: The reserve requirements in Table 2 are pooled across the Southeast region. Pooled reserves are calculated for the four balancing regions as a whole by using the sum of solar and wind profiles for the four balancing regions to calculate solar and wind forecast errors. We include wind resources in the reserve calculations for completeness, but most of the solar and wind forecast error is driven by solar resources because solar installed capacity (137 GW) is much larger than wind (6 GW).

2.4 Production Simulation

We use two-stage production simulation modeling to mimic day-ahead scheduling and real-time dispatch processes. All production simulation in this analysis uses the PLEXOS energy modeling software. We simulate the day-ahead and real-time stages separately. Figure 2 shows the structure for each stage. For the day-ahead stage, the model solves each day sequentially, passing each unit's commitment status (on or off), up and down times, schedule, and battery state-of-charge (SOC) to the next day. Each day's forecast includes at least a one-day lookahead. For instance, with a two-day day-ahead lookahead (two days after the next day), the day-ahead simulation for January 2 uses 24-hour-ahead (January 2), 48-hour-ahead (January 3), and 72-hour-ahead (January 4) forecasts that would have all been generated on January 1. The model solves for the entire period (in this case, January 2 to 4) but only keeps the results of the day being simulated (January 2).

Day-ahead scheduling simulations



Real-time dispatch simulations

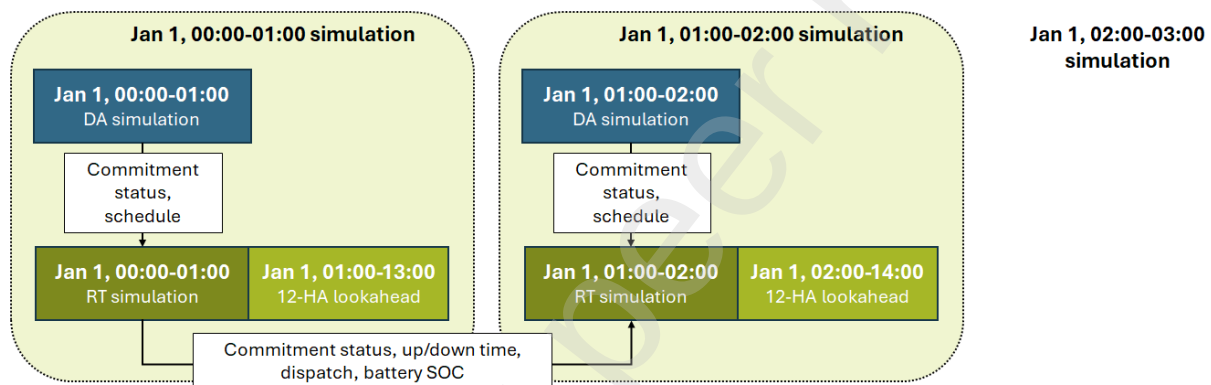


Figure 2. Illustration of day-ahead and real-time simulations

Once the day-ahead simulations are complete, the model solves 15-minute real-time dispatch for each hour sequentially. The model passes day-ahead commitment and dispatch results for that hour and commitment, up and down time, dispatch, and battery SOC from the previous hour to that hour's real-time simulation. Each hourly real-time simulation contains a multi-hour lookahead. For instance, a 12-hour lookahead from hour beginning (H.B.) 00:00 means that the model solves for the period from 00:00 to 13:00 but only keeps the results of the dispatch interval (00:00-01:00).

The model allows users to specify both the time horizon and resolution of lookahead periods for day-ahead scheduling and real-time dispatch. For solar forecasts, both the forecast horizon and its resolution impact operations [27,28]. In this analysis, we use base lookahead configurations and then consider longer day-ahead and real-time lookahead periods as sensitivities. In the base case day-ahead simulations, we use a 24-hour day-ahead lookahead (one day lookahead after day-ahead) with 4-hour time resolution. In the base case real-time simulations, we use a 12-hour lookahead (12-hour lookahead after the real-time interval) with 4-hour time resolution. In sensitivity cases, we alternately examine a 48-hour day-ahead lookahead with 4-hour time resolution and a 24-hour real-time lookahead with 4-hour time resolution. All of the day-ahead lookaheads use solar forecasts (NAM or persistence) rather than solar profile data. The real-time lookaheads use solar profile data rather than solar forecasts,

assuming perfect foresight. Load and wind use profile data for both day ahead and real time, thus assuming perfect foresight rather than forecasts.

Figure 3 illustrates the time horizon and resolution of lookaheads for solar. For day D (January 1 in the figure), the day-ahead solar forecast is hourly and covers January 2 (D+1). For day D+2 (24-hour lookahead), the day D solar forecast covers January 3 in 4-hour blocks. In the longer day-ahead lookahead sensitivity, the day D solar forecast covers January 4 (day D+3, 48-hour lookahead) in 4-hour blocks. In real-time, each hour has 4 15-minute intervals with actual available solar data and, in the base case, uses a 12-hour lookahead in 4-hour blocks. The longer real-time lookahead sensitivity adds an additional 12 hours to the real-time lookahead, in 4-hour blocks.

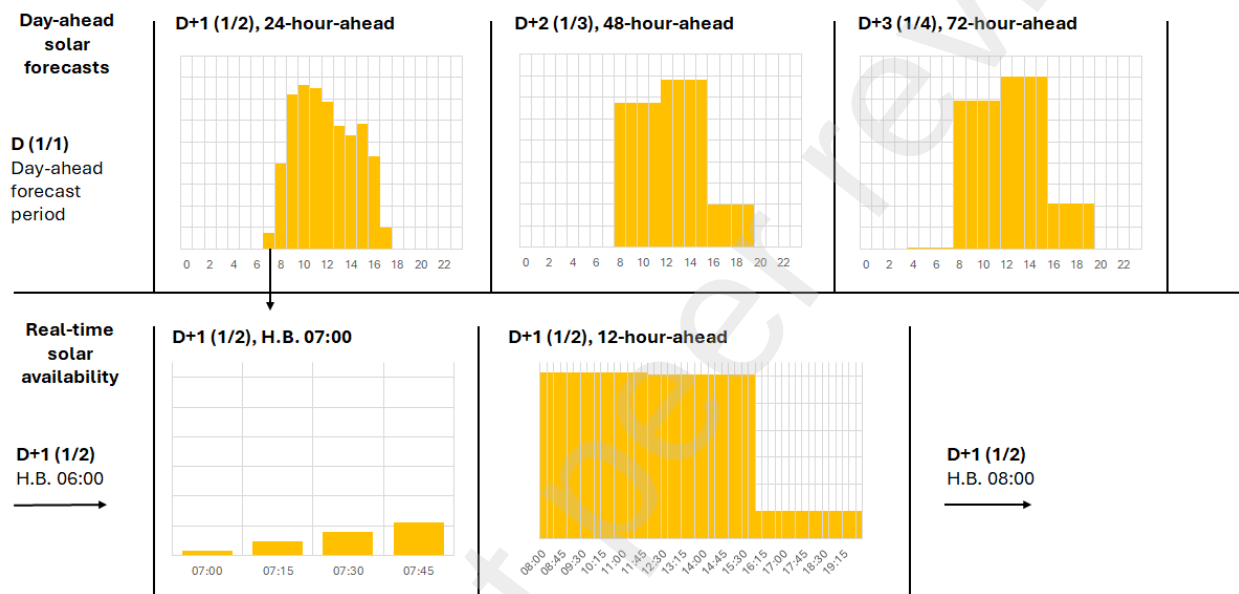


Figure 3. Illustration of solar forecasts and profiles in day-ahead and real-time lookaheads

Different types of generation units have different limits on the extent that they can be turned on and off and re-dispatched in real-time. For nuclear, hydro, and distributed PV generation, we do not allow real-time re-dispatch of day-ahead schedules. For coal, gas combined cycle (GCC), biomass, and other steam generation, we do not allow real-time commitment (or de-commitment) but allow real-time re-dispatch. In a flexible gas scenario, we allow GCC generation to be committed and de-committed in real-time. For simple cycle gas combustion turbines (GCT), wind, utility-scale PV and energy storage, we allow both commitment/de-commitment and re-dispatch in real-time.

We limit the production simulation modeling to the four balancing regions in the Southeast. We consider two scenarios for how the four balancing regions interact. In a low coordination scenario, we assume relatively high (\$32/MWh, 2023\$) hurdle rates among balancing regions and require all balancing regions to hold their own operating reserves, though we allow forecast error reserves to be pooled across regions. In a high coordination scenario, we remove hurdle rates (\$0/MWh), allow forecast error reserve sharing, and allow all operating reserves to be held in any of the four balancing regions.

2.5 Modeling Scenarios

The analysis considers three scenarios and two sensitivities (Table 3). The Base scenario uses a 24-hour day-ahead lookahead, a 12-hour real-time lookahead, does not allow GCC units to be turned on or off in real-time, imposes a \$32/MWh hurdle rate between regions, and requires that each balancing region hold its own reserves. The High Coordination (HighCoord) scenario is the same as the Base scenario except that it removes hurdle rates between regions and optimizes reserve location throughout the Southeast region. The Flexible Gas (FlexGas) scenario is the same as the Base scenario except that allows GCC units to be turned on and off in real-time dispatch. The Longer Day-Ahead Lookahead (LongDALook) sensitivity is the same as the Base scenario but uses a longer day-ahead lookahead. We also examine a Longer Real-time Lookahead (LongRTLook) sensitivity in Appendix B.4.

Table 3. Scenarios and Sensitivities Examined in this Analysis

Scenario	
Base	• 24-hour day-ahead lookahead • 12-hour real-time lookahead • Fixed day-ahead GCC commitment • \$32/MWh hurdle rates • Balancing regions hold reserves
HighCoord	Same as Base except: • \$0/MWh hurdle rates • Optimized reserve location
FlexGas	Same as Base except: • RT GCC commitment
Sensitivity	
LongDALook	Same as Base except: • 48-hour day-ahead lookahead
LongRTLook	Same as Base except: • 24-hour real-time lookahead

3. Results

This section is organized into six sub-sections by topic area: (1) production costs, (2) unit commitment and dispatch, (3) emissions, (4) operations during days with the worst forecast errors, (5) reserve violations and capacity needs, and (6) the day-ahead lookahead sensitivity. In most cases, we report results for the four balancing regions as a whole rather than for each balancing region individually.

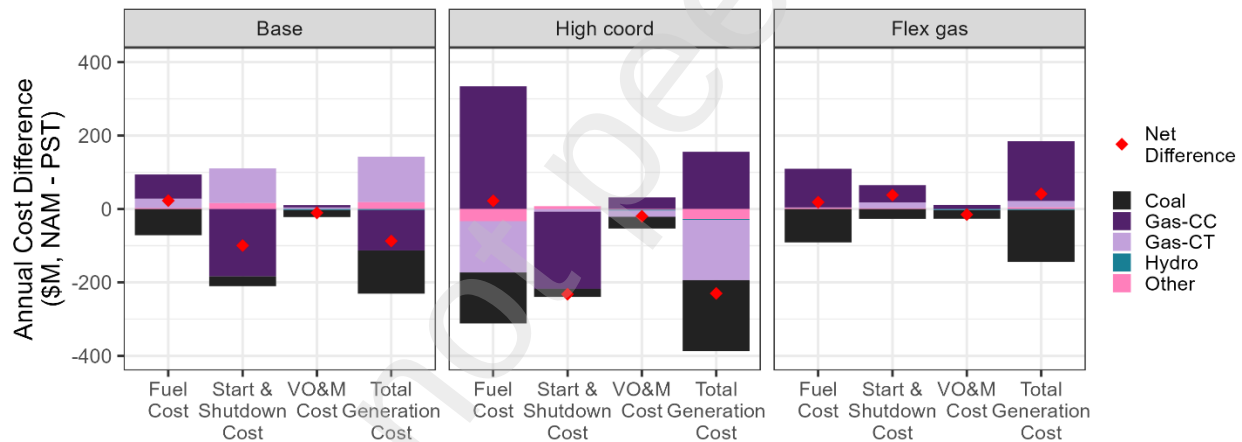
3.1 Production Costs

Table 4 shows production costs by component, total costs (sum of component costs), and total unit costs (cost divided by annual load), for each of the three scenarios and the NAM and persistence (PST) forecast cases. Improved forecast accuracy in the NAM case reduces production costs by \$87 million per year (\$0.13/MWh [2023\$], or by 0.9%) in the Base scenario and \$230 million per year (\$0.34/MWh, or by 2.5%) in the HighCoord scenario (more import and export flexibility), but increases production costs by \$41 million per year (\$0.06/MWh, or by 0.5%) in the FlexGas scenario (higher GCC flexibility).

Table 4. Production Costs (2023\$) by Scenario for NAM and Persistence (PST) Forecast Cases

	Base		HighCoord		FlexGas	
	NAM	PST	NAM	PST	NAM	PST
Start costs (M\$)	\$1,223	\$1,323	\$1,437	\$1,669	\$1,335	\$1,297
Fuel costs (M\$)	\$6,999	\$6,976	\$6,687	\$6,665	\$6,593	\$6,575
VO&M costs (M\$)	\$901	\$911	\$794	\$814	\$863	\$878
Total costs (M\$)	\$9,124	\$9,211	\$8,918	\$9,148	\$8,791	\$8,750
Total unit costs (\$/MWh)	\$13.32	\$13.45	\$13.02	\$13.36	\$12.83	\$12.77

Figure 4 shows total production cost changes (NAM minus PST) by production cost component, generation type, and scenario. In the Base scenario, cost savings are mainly due to lower GCC start costs and coal fuel costs, which outweigh the increase in GCC fuel and GCT start and fuel costs. In the HighCoord scenario, cost savings are mainly due to lower GCC start costs and coal and GCT fuel costs, which outweigh the increase in GCC fuel costs. In the FlexGas scenario, cost increases are mainly due to higher GCC fuel and start costs, which outweigh the decrease in coal fuel and start costs. The overall cost increase in the FlexGas scenario is likely due to sub-optimal coal and gas commitment with the NAM forecast (Appendix B.2).



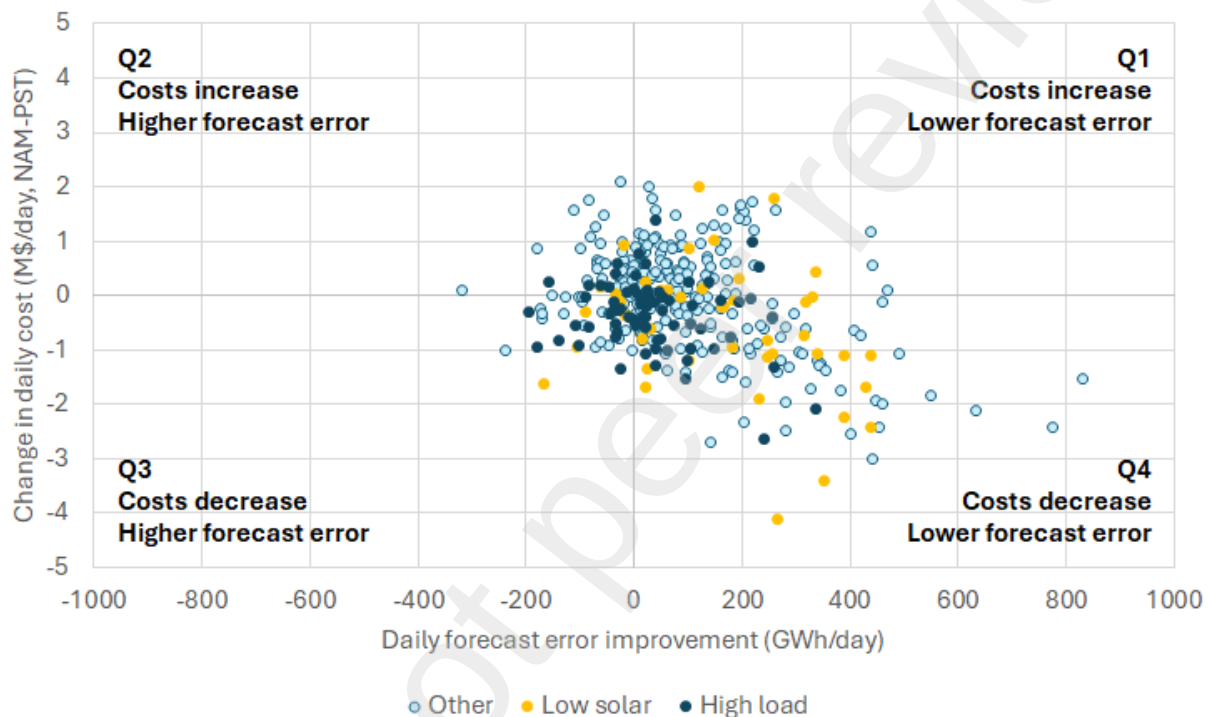
Notes: Other includes oil and gas steam units and biomass units.

Figure 4. Changes in production costs in NAM relative to PST (NAM minus PST) by generation type and scenario

The changes in costs shown in Figure 4 have complex drivers, due to the multi-day horizon for startup and shutdown decision-making and the structural differences between the NAM and PST forecasts. To explore cost drivers, we performed a simple regression of daily cost changes (NAM – PST) on daily solar availability, load levels, forecast improvements, and seasons using the Base scenario (see Appendix B.1 for a more detailed description). Daily cost changes are significantly correlated with daily solar availability (higher cost savings on low solar days [$p < 0.01$]), load levels (higher savings on high load days [$p < 0.01$]), PST forecast error (higher savings with higher absolute errors [$p < 0.01$]), forecast improvements (higher savings with higher forecast improvements [$p < 0.01$]), and with some seasons

(higher savings in spring [$p < 0.05$] and summer [$p < 0.01$]). However, these variables explain only a limited amount of the variability in daily cost changes (adjusted $R^2 = 0.29$).

Beyond correlations, it is difficult to find clearer patterns and draw more general conclusions about the relationship between production cost changes and solar forecast characteristics. Figure 5 shows changes in daily production costs between the NAM and PST forecasts for low solar days, high load days, and all other days as a function of forecast improvements in NAM for the Base scenario. Daily forecast improvements generally, but not always, reduce daily costs. Larger changes in daily costs do not disproportionately occur on low solar or high load days relative to other days.



Notes: Q1-Q4 indicate the four quadrants. Cost increase means that daily costs in the NAM case are higher than daily costs in the PST case. Higher forecast error means that daily forecast error in the NAM case is higher than daily forecast error in the PST case. Low solar and high load days are defined using an 80% exceedance probability. Daily forecast error improvement is the reduction in the absolute forecast error between the NAM and PST cases.

Figure 5. Change in daily production costs in the Base scenario as a function of daily forecast error improvement (NAM-PST) for low solar days, high load days, and other days

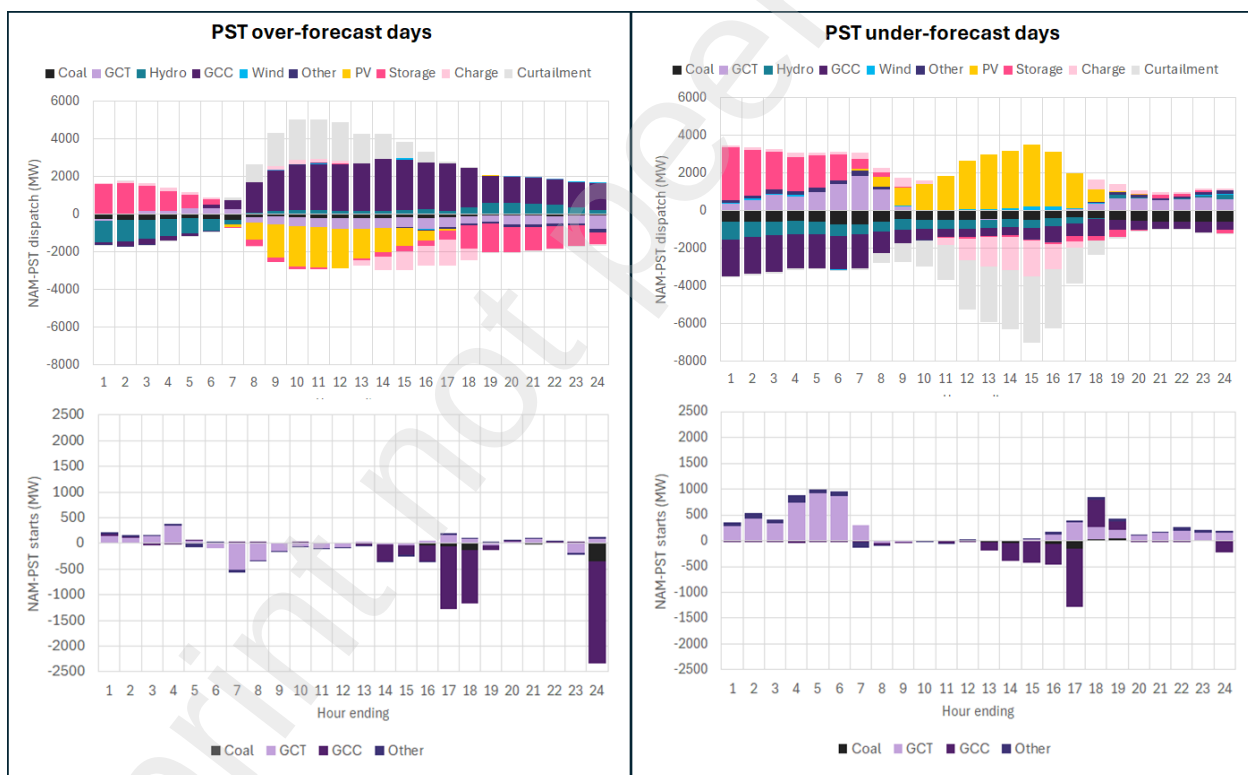
3.2 Unit Commitment and Dispatch

We define over-forecast days as days in which actual daily solar (MWh per day) is less than forecasted daily solar, and under-forecast days as days in which actual daily solar is greater than forecasted daily solar. Appendix B.2 describes unit commitment and dispatch results for the HighCoord and FlexGas scenarios

Figure 6 shows changes in real-time average daily dispatch (top panels) and day-ahead and real-time average daily starts (total MW, bottom panels) between the NAM and PST forecasts in the Base scenario. On days with a PST over-forecast (actual solar < forecast, lefthand side of Figure 6), in the NAM

case the model de-commits less GCC capacity day-ahead than the PST case during the middle of the day and commits less GCC capacity from 14:00 to 18:00 (hour ending). As a result, in the NAM case the model has more GCC capacity available to manage lower than expected solar availability but curtails more solar due to higher total minimum generation levels. With less GCC generation available during the real-time evening solar ramp, in the PST case the model discharges more storage, reduces storage charging, and dispatches more GCT units than in the NAM case in the mid- to late-afternoon. In the NAM case, the model starts less GCC capacity in hour 24:00 than in the PST case, which is difficult to attribute to specific forecast events.

On days with a PST under-forecast (actual solar > forecast), in the NAM case the model absorbs and stores more solar generation (less curtailment) than in the PST case, discharges more storage from 01:00-07:00 due to higher solar generation, and starts and dispatches more GCT generation in the evenings and mornings. As a result, the NAM case has less GCC, coal, and hydro generation and commits less GCC capacity than in the PST case. It is important to keep in mind that the differences in GCC starts all occur on a day-ahead time horizon, because GCC cannot be committed or de-committed in real-time in the Base scenario, and thus should be seen as a day-ahead response to different forecasts rather than a real-time response to forecast errors.



Notes: A positive (negative) value for generation, storage discharge, and curtailment means that they are higher (lower) in the NAM case than in the PST case. A positive (negative) value for storage charging means that it is lower (higher) in the NAM case than in the PST case. If storage losses (not shown) were included, the hourly differences between the NAM and PST cases would exactly balance. Curtailment is shown though it is not part of the NAM minus PST energy balance because the figure shows actual solar generation rather than available solar energy.

Figure 6 Error! Reference source not found.. Changes in average daily dispatch (top plots) and starts (bottom plots) between the NAM and PST forecasts for PST over-forecast and under-forecast days, Base scenario

Table 5 shows changes in annual real-time generation, storage charging, and curtailment between the NAM and PST cases for the Base scenario. The largest changes are in GCC generation (+3.1 TWh in NAM relative to PST) and coal generation (-3.0 TWh). As illustrated in Figure 6, changes in coal and GCC generation are the result of interactions involving solar, storage, GCT, and hydropower generation rather than direct substitution.

Table 5. Annual Generation (TWh/yr) in the NAM and PST Cases, Change in Annual Generation (Δ , NAM-PST), and Percentage Change ($\% \Delta$) in Annual Generation, Base Scenario

	Coal	GCC	GCT	OTH	NUC	HYD	WND	PV	STD	STC	CURT
NAM	59	117	10	5	216	32	22	237	66	77	68
PST	62	114	9	5	216	33	22	236	65	76	69
Δ	-3.0	3.1	0.6	0.1	0.0	-1.6	0.2	0.9	1.2	1.5	-1.1
$\% \Delta$	-5%	3%	6%	2%	0%	-5%	1%	0%	2%	2%	-2%

Notes: OTH is other thermal, including oil and gas steam and biomass; NUC is nuclear; HYD is hydropower, WND is onshore and offshore wind; PV includes distributed PV; STD and STC are storage discharge and charge, respectively; CURT is curtailment.

3.3 Emissions

NAM carbon dioxide (CO_2) emissions are lower than PST emissions in all three scenarios (Table 6). Lower emissions are primarily due to coal to natural gas switching, as described in the previous two sections. Emission reductions were larger than expected. To provide a sense of relative scale, production cost savings in the Base scenario are equivalent to \$58/t CO_2 of emission reductions.

Table 6. NAM and PST CO_2 Emissions by Scenario (Mt CO_2)

Scenario	NAM	PST	NAM – PST
Base	112.5	114.0	-1.5
HighCoord	99.2	102.9	-3.7
FlexGas	108.3	110.5	-2.2

3.4 Operations During Days with the Worst Forecasts

Figures 7 and 8 show average day-ahead and real-time dispatch, solar availability, and storage charging in the Base scenario during the 10% of days with the worst PST daily under- and over-forecasts. The top panels show daily average generation, whereas the bottom panels show the difference between the NAM and PST cases. In Figure 7 (under-forecast), real-time solar availability in the PST case is significantly larger than that in day-ahead. In the PST case, the model addresses solar forecast error primarily by increasing solar curtailment, increasing storage charge and discharge, and reducing GCC generation. The NAM forecast is significantly more accurate on average during these days. The main difference between PST and NAM (bottom-most panel) in day-ahead is in curtailment and storage charge and discharge. Real-time differences between NAM and PST are relatively small, despite the large difference in solar forecast errors.

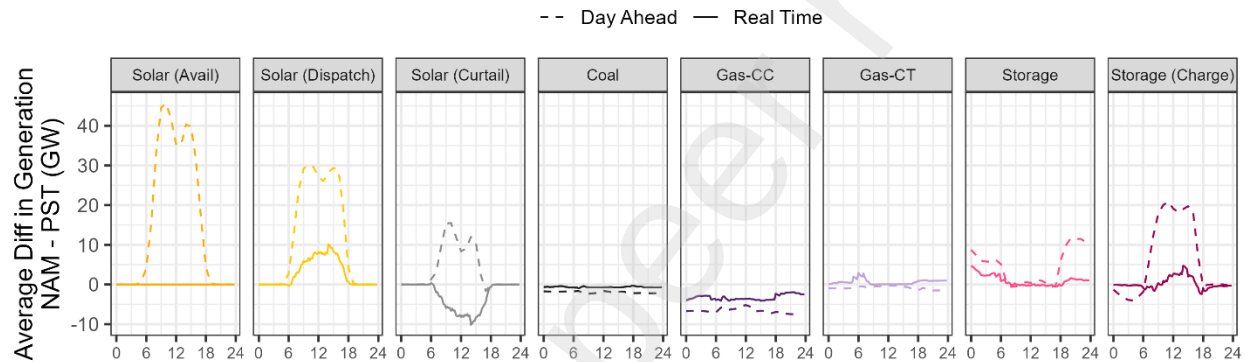
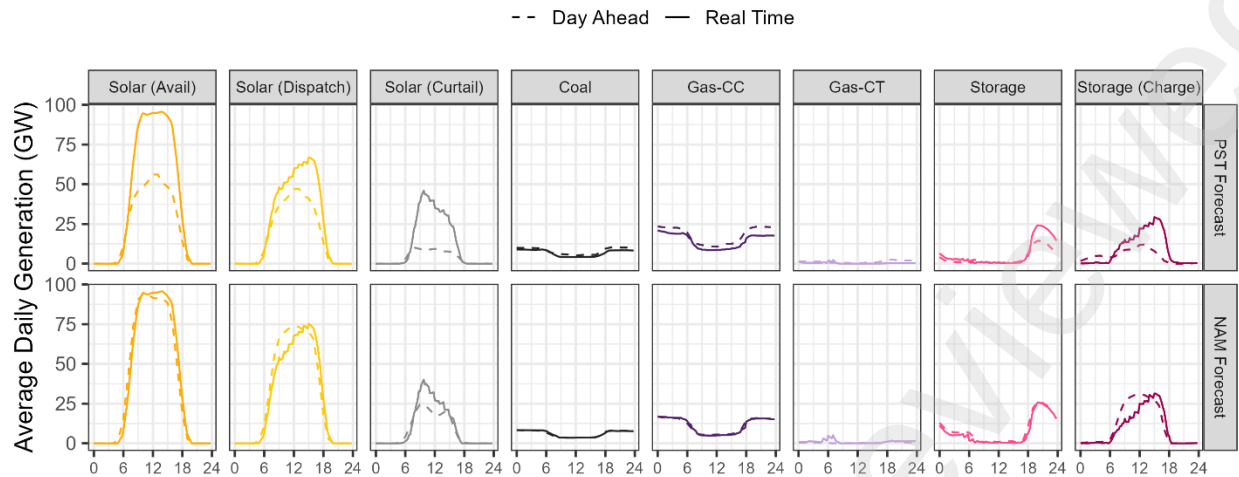


Figure 7. PST and NAM average daily generation, solar availability, and storage charging in the days with the 100 worst under-forecast hours, and differences between PST and NAM

In Figure 8 (over-forecast), in the PST case solar availability is significantly lower in real-time than in day-ahead. In the PST case, the model addresses solar forecast error primarily by reducing day-ahead scheduled curtailment and, to a lesser extent, by reducing storage charging and increasing coal and GCC generation in the middle of the day and GCT generation throughout the day. The NAM forecast is more accurate on average during these days. Day-ahead, the main difference between the PST and NAM cases is in curtailment and storage charging. Real-time differences are small relative to day-ahead differences.

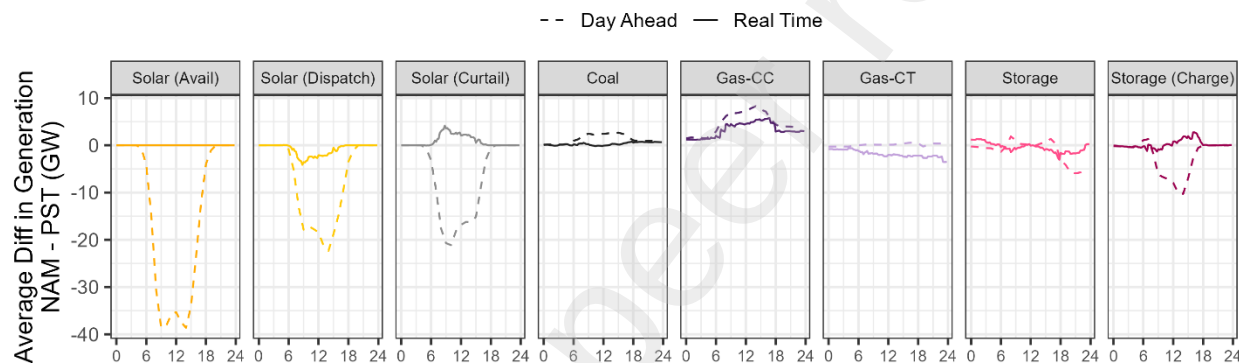
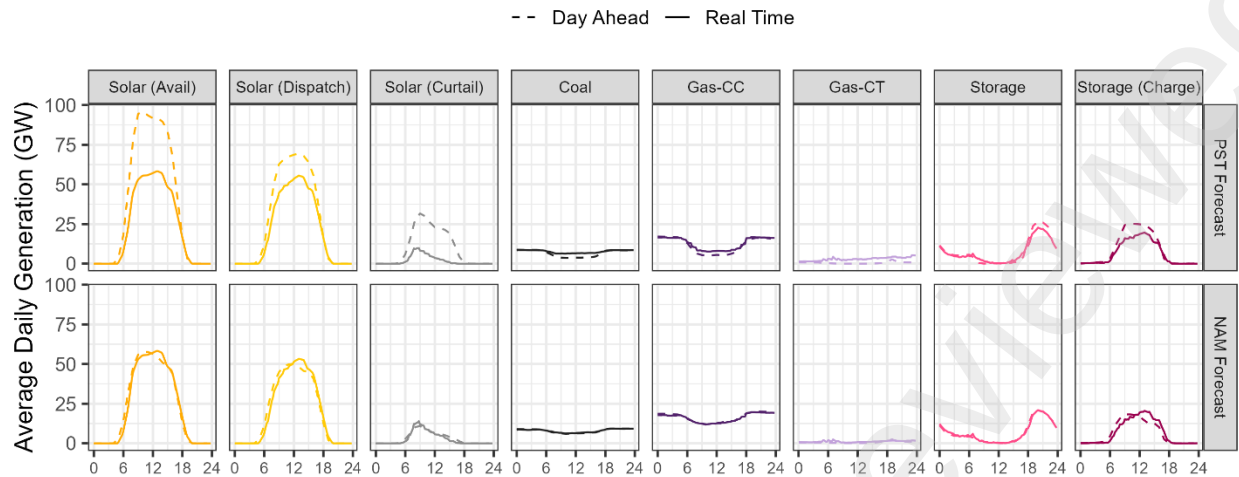


Figure 8. PST and NAM average daily generation, solar availability, and storage charging in the fays with the 100 worst over-forecast hours, and differences between PST and NAM

These results illustrate the important role that curtailment plays in responding to solar forecast error in high-solar systems. Across scenarios, real-time curtailment ranged from 15% to 20% of available solar, wind, and hydro generation. In the Base scenario, curtailment occurred in 66% and 69% of hours in the NAM and PST cases, respectively, and 32%-39% in the HighCoord and FlexGas scenarios. When curtailment is already scheduled day-ahead and there is an over-forecast, higher real-time curtailment can help to manage forecast error (Figure 7). When there is an under-forecast, reduced real-time curtailment can help to manage forecast error (Figure 8).

3.5 Reserve Violations and Capacity Needs

Table 7 shows day-ahead reserve violations for the NAM and PST cases in the Base scenario. Due to large day-ahead non-spinning forecast error reserve requirements (see Section 2.3), day-ahead reserve violations were frequent and large in the PST and, to a lesser extent, in the NAM case. In both cases, most day-ahead reserve violations were in non-spinning reserves. Real-time reserve violations were small in both NAM and PST, with no significant difference between them (see Appendix B.3).

376 **Table 7. Day-Ahead Reserve Violations in NAM and PST, Base Scenario**

	Forecast error spinning	Forecast error non- spinning	Contin- gency spinning	Contin- gency non- spinning	Regulation
NAM					
Frequency (number of hours, % annual hours)	42 (0%)	355 (4%)	19 (0%)	533 (6%)	1 (0%)
Total violations (GW-h/yr)	38	2,397	12	335	0
Max violation (GW)	9	27	4	4	0
Average violation (MW-h/event, % reserves)	903 (10%)	6,753 (27%)	626 (17%)	628 (18%)	476 (7%)
PST					
Frequency (number of hours, % annual hours)	266 (3%)	1,543 (18%)	157 (2%)	1,696 (19%)	0 (0%)
Total violations (GW-h/yr)	84	13,848	29	758	0
Max violation (GW)	12	53	4	4	0
Average violation (MW-h/event, % reserves)	314 (2%)	8,975 (21%)	187 (6%)	447 (15%)	0 (0%)

377 Notes: Total amount (in GW-h per year) is the sum of all reserve violations during the year. Average lost is the
378 average amount of reserves lost per violation event, in absolute (MW per hour or MW-h) and relative (share of
379 reserves) terms.

380 3.6 Extended Day-Ahead Lookahead

381 The use of a 48-hour rather than a 24-hour lookahead period in the longer day-ahead look ahead
382 (LongDALook) sensitivity results in slightly lower production cost savings with the improved NAM
383 forecast (\$80 million per year in savings) compared to the Base scenario (\$87 million per year in savings
384 [Table 8]). The dynamics behind this counterintuitive result can be seen in the production cost results.
385 The longer day-ahead lookahead lowers costs in both the NAM and PST case, relative to the Base
386 scenario, but start cost reductions are larger in the NAM case, fuel costs reductions are larger in the PST
387 case, and total cost reductions are larger in the PST case. Thus, the NAM-PST difference between costs
388 in the LongDALook is lower than in the Base scenario, though the difference is small.

389 **Table 8. Annual Production Costs (Million 2023\$) and CO₂ Emissions (MtCO₂) in the Base Scenario and**
390 **LongDALook Sensitivity**

	Base		LongDALook	
	NAM	PST	NAM	PST
Start costs (M\$)	\$1,223	\$1,323	\$1,198	\$1,320
Fuel costs (M\$)	\$6,999	\$6,976	\$6,977	\$6,925
Variable O&M costs (M\$)	\$901	\$911	\$901	\$911
Total costs (M\$)	\$9,124	\$9,211	\$9,075	\$9,155
CO₂ emissions (MtCO₂)	112.5	114.0	113.0	114.2

391
392 The longer lookahead in the LongDALook sensitivity leads to negligible changes in real-time solar and
393 storage utilization relative to the Base scenario. In both NAM and PST, differences in annual storage
394 charge and discharge, solar generation, curtailment, and average hourly storage SOC are all within 1%.

Differences in hourly average storage discharge between the LookDALook sensitivity and Base scenario are relatively small (within 8%). This implies that, at these levels of storage, solar, GCC, and GCT capacity, multi-days-ahead solar forecasts have a limited impact on storage utilization and optimization.

The cost savings from a longer day-ahead lookahead with higher solar forecast accuracy (LongDALook NAM versus LongDALook PST, with \$80 million in cost savings per year) are higher than the cost savings from a longer lookahead with the same level of forecast accuracy (LongDALook NAM versus Base NAM, with \$49 million in cost savings per year; LongDALook PST versus BaseLook PST, with \$55 million in cost savings per year), suggesting that the accuracy of the lookahead period does significantly impact the results.

4. Conclusions

In the Base scenario, annual production cost savings from higher solar forecast accuracy are \$87 million per year (\$0.13/MWh [2023\$], or a 0.9% reduction in costs). These cost savings are within the range of savings in the Hodge et al. [9] and Hobbs et al. [11] studies (\$0.04/MWh-\$0.28/MWh) and lower than in the Martínez-Anido [7,8] study (\$0.11-\$1.42/MWh) (see Section 1). In earlier studies, solar generation was less than 25% of total generation, versus 34%-36% across scenarios in this study. At the higher levels of solar and storage in this study, the value of improved solar forecasting did not increase in absolute terms relative to studies with less solar and storage.

Higher system flexibility can increase or decrease the value of forecast accuracy depending on the type of system flexibility. With more regional flexibility (lower hurdle rates, optimized reserves) in the HighCoord scenario, production cost savings from solar forecast improvements are \$230 million per year (\$0.34/MWh, 2.5% reduction in costs relative to the Base scenario). However, with more GCC commitment flexibility (real-time gas commitment), higher solar forecast accuracy increases production costs by \$41 million per year (\$0.06/MWh, 0.5% increase in costs).

Solar curtailment and energy storage play a central role in managing day-ahead solar forecast error. On days with the worst persistence forecast errors, the model manages forecast error-driven imbalances between day-ahead and real-time mainly through changes in day-ahead solar curtailment and storage charge and discharge operations. Changes in day-ahead curtailment are an important strategy for managing solar forecast error because solar curtailment occurs frequently. In the Base scenario, curtailment occurs in 66% (NAM) and 69% (PST) of hours. If solar capacity expansion is accompanied by higher levels of energy storage and more frequent curtailment, as in this study, solar development will bring its own source of flexibility for managing solar forecast error.

Because of the flexibility provided by solar curtailment and storage, differences in real-time dispatch and costs between the NAM and persistence forecasts are mostly in thermal generation. On an annual basis, the main differences are in GCC (3%-11% absolute difference across scenarios), coal (5%-11%), and GCT (0%-29%) generation rather than in solar generation (0%-1%), storage charge and discharge (0%-2%), and curtailment (2%). GCC starts are the most important driver of production cost differences between solar forecasts and between scenarios.

Lower reserve capacity to manage solar forecast error is an often overlooked but potentially significant benefit of improved solar forecasts. In this study, the NAM forecast reduces reserves needed to manage solar forecast errors by 30%. With a less accurate forecast, reserves needed to manage two standard deviations of infrequent, large solar forecast errors may exceed planning reserves. If so, an improved

solar forecast may reduce the capacity cost of reserves to manage forecast error, a benefit that could be larger than production cost savings. The capacity value of improved solar forecasts will be specific to system operators, depending on their risk tolerance, and would ideally be assessed through capacity expansion models and resource adequacy programs.

5. Acknowledgements

Funding for this study was provided by the U.S. Department of Energy's (DOE's) Office of Energy Efficiency and Renewable Energy Strategic Analysis, Solar Energy Technologies Office. The study was authored in part by Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231, and in part by the National Renewable Energy Laboratory for the DOE under Contract No. DE-AC36-08GO28308. This research was performed using computational resources sponsored by the DOE's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory. We especially thank Ammar Qusaibaty of DOE for supporting this work. We also acknowledge helpful support from Kate Doubleday, Clayton Barrows, Kate Anderson, and Jaquelin Cochran.

The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

6. References

- [1] E.W. Law, M. Kay, R.A. Taylor, Calculating the financial value of a concentrated solar thermal plant operated using direct normal irradiance forecasts, *Solar Energy*. 125 (2016) 267–281.
- [2] J. Luoma, P. Mathiesen, J. Kleissl, Forecast value considering energy pricing in California, *Applied Energy*. 125 (2014) 230-237.
- [3] A. Kaur, L. Nonnenmacher, H.T.C. Pedro, C.F.M. Coimbra, Benefits of solar forecasting for energy imbalance markets, *Renewable Energy*. 86 (2016) 819-830.
- [4] B. Kraas, M. Schroedter-Homscheidt, R. Madlener, Economic merits of a state-of-the-art concentrating solar power forecasting system for participation in the Spanish electricity market, *Solar Energy*. 93 (2013) 244-255.
- [5] J. Antonanzas, D. Pozo-Vázquez, L.A. Fernandez-Jimenez, F.J. Martinez-de-Pison, The value of day-ahead forecasting for photovoltaics in the Spanish electricity market, *Solar Energy*. 158 (2017) 140-146.
- [6] Y. Wang, D. Millstein, A.D. Mills, S. Jeong, A. Ancell, The cost of day-ahead solar forecasting errors in the United States, *Solar Energy*. 231 (2022) 846-856.
- [7] C.B. Martínez-Anido, A. Florita, B-M. Hodge, The Impact of Improved Solar Forecasts on Bulk Power System Operations in ISO-NE, NREL Technical Report NREL/CP-5D00-62817, Golden, CO, NREL, 2014.
- [8] C.B. Martínez-Anido, B. Botor, A.R. Florita, C. Draxl, S. Lu, H.F. Hamann, B-M. Hodge, The value of day-ahead solar power forecasting improvement, *Solar Energy*. 129 (2016) 192-203.

472 [9] B-M. Hodge, C. Brancucci Martinez-Anido, Q. Wang, E. Chartan, A. Florita, J. Kiviluoma, The combined
473 value of wind and solar power forecasting improvements and electricity storage, *Applied Energy*. 214
474 (2018) 1-15.

475 [10] S. Goodarzi, H.N. Pererac, D. Bunne, The impact of renewable energy forecast errors on imbalance
476 volumes and electricity spot prices, *Energy Policy*. 134 (2019) 110827.

477 [11] W.B. Hobbs, J. Tresan, M. Kline, M. Guha, B. Duncan, Modeled Impacts of Solar Forecast Error on
478 Utility Production Cost, *TechRxiv*. 2023.

479 [12] M. Cui, J. Zhang, B-M., Hodge, S. Lu, H.F., Hamann, A Methodology for Quantifying Reliability
480 Benefits From Improved Solar Power Forecasting in Multi-Timescale Power System Operations, *IEEE*
481 *Transactions on the Smart Grid*. 9 (2018) 6897-6908.

482 [13] O. Ghandi, W. Zhang, D.S. Kumar, C.D. Rodriguez-Gallegos, G.M. Yagli, D. Yang, T. Reindl, D.
483 Srinivasan, The value of solar forecast and the cost of their errors: A review, *Renewable and Sustainable*
484 *Energy Reviews*. 189 (2024) 113915.

485 [14] K. Porter, J. Rogers. Survey of Variable Generation Forecasting in the West, NREL/SR-5500-54457.
486 Golden, CO, NREL, 2012.

487 [15] California Independent System Operator (CAISO), GMC and Other Rates for 2004-2023, 2023.

488 [16] CAISO, 2023 Annual Report on Market Issues & Performance, Folsom, CA, CAISO, 2024.

489 [17] F. Kahrl, J. Jorgenson, L. Kiboma, J.H. Kim, D. Millstein, B. Sergi, N.M. Frick. Solar and Storage
490 Integration in the Southeastern United States Economics, Reliability, and Operations, Berkeley, CA,
491 LBNL, 2024.

492 [18] J. Ho, J. Becker, M. Brown, P. Brown, I. Chernyakhovskiy, S. Cohen, W. Cole, S. Corcoran, K. Eureka,
493 W. Frazier, P. Gagnon, N. Gates, D. Greer, P. Jadun, S. Khanal, S. Machen, M. Macmillan, T. Mai, M.
494 Mowers, C. Murphy, A. Rose, A. Schleifer, B. Sergi, D. Steinberg, Y. Sun, E. Zhou, Regional Energy
495 Deployment System (ReEDS) Model Documentation: Version 2020, NREL Technical Report NREL/TP-
496 6A20-78195, Golden, CO, NREL, 2021.

497 [19] CAISO, Market Performance and Planning Forum, December 15, 2020.

498 [20] CAISO, Market Performance and Planning Forum, June 29, 2023.

499 [21] J. Zhang, A. Florita, B. Hodge, S. Lu, H.F. Hamann, V. Banunarayanan, A.M. Brockway, A suite of
500 metrics for assessing the performance of solar power forecasting, *Solar Energy*. 111 (2015) 157-175.

501 [22] E. Ela, V. Diakov, E. Ibanez, M. Heaney, Impacts of Variability and Uncertainty in Solar Photovoltaic
502 Generation at Multiple Timescales, NREL Technical Report NREL/TP-5500-58274, Golden, CO, NREL,
503 2013.

504 [23] GE Energy, Western Wind and Solar Integration Study, NREL Subcontract Report NREL/SR-550-
505 47434, Golden, CO, NREL, 2010.

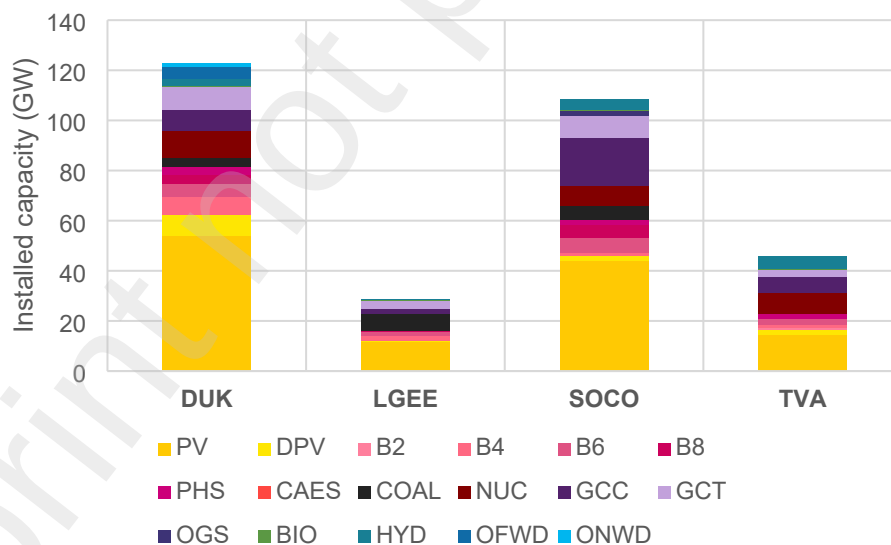
506 [24] EnerNex, Eastern Wind Integration and Transmission Study, NREL Subcontract Report NREL/SR-
507 5500-47078, Golden, CO, NREL, 2011.

- [25] D. Lew, G. Brinkman, E. Ibanez, A. Florita, M. Heaney, B.-M. Hodge, M. Hummon, G. Stark, J. King, S.A. Lefton, N. Kumar, D. Agan, G. Jordan, S. Venkataraman, The Western Wind and Solar Integration Study Phase 2, NREL/TP-5500-55588, Golden, CO, NREL, 2013.
- [26] X. Fang, J. Tan, H. Yuan, S. Yin, J. Wang, Providing Ancillary Services with Photovoltaic Generation in Multi-Timescale Grid Operation, NREL/CP-5D00-75321, Golden, CO, NREL, 2021.
- [27] A. Tuohy, J. Zack, S.E. Haupt, J. Sharp, M. Ahlstrom, S. Dise, Solar Forecasting: Methods, Challenges, and Performance, IEEE Power and Energy Magazine. 13 (2015) 50-59.
- [28] K. Doubleday, V. Van Scyoc Hernandez, B. Hodge, Benchmark probabilistic solar forecasts: Characteristics and recommendations, Solar Energy. 206 (2020) 52-67.

Appendix A. Additional Methods

A.1. Medium PV High Storage Scenario Resource Portfolios

Figure A.1.1 shows the medium PV high storage (MPHS) resource portfolios for each of the four balancing regions. The MPHS portfolios are from Kahrl et al. [17] and are for the year 2035. In total, the four regions have 137 GW of installed solar photovoltaic (PV) capacity (45% of total installed capacity) and 43 GW / 292 GWh of energy storage (33% of coincident peak demand, 0.04% of annual energy consumption). More than 70% of installed storage capacity is 6-hour batteries (36%), 8-hour batteries (21%), and 12-hour pumped hydro storage (15%). Three of the four balancing regions (DUK, TVA, SOCO) have substantial nuclear (27 GW total) and hydropower (13 GW) generation, and three regions (LGEE, SOCO, DUK) have substantial coal (17 GW) generation. All four regions have significant amounts of simple cycle (24 GW) and combined cycle (36 GW) natural gas generation.



Notes: DPV is distributed PV; B2, B4, B6, and B8 are battery storage with 2, 4, 6, and 8-hour duration; PHS is pumped hydro storage; CAES is compressed air energy storage; NUC is nuclear; GCC is natural gas combined cycle; GCT is natural gas combustion turbine; OGS is oil and gas steam; BIO is biomass; HYD is hydropower; OFWD and ONWD are offshore and onshore wind.

Figure A.1.1. Installed capacity by balancing region for the high-solar, high-storage scenario used in this study

A.2. Solar Data Development

We use the North American Mesoscale (NAM) forecast system to develop day-ahead forecasts and real-time profile data. We use the nodes in the ReEDS model as locations for NAM forecasts and profiles. NAM predicts weather and solar irradiance for up to 84 hours ahead at 12 km spatial resolution and 1-hour temporal resolution. NAM is run 4 times each day at 00:00, 06:00, 12:00, and 18:00 coordinated universal time (UTC). Each NAM run is initialized with a six hour spin-up period, during which measured meteorological data is assimilated into the model before running the forecast. We combine the six closest forecast values (zero to five hours ahead) from each NAM run into a continuous time series for real-time solar profiles. For instance, the 06:00 UTC cycle output from January 2, 2012, was used to obtain the six data points corresponding to the 06:00-11:00 UTC hours for that day. Similarly, we combined the forecasts from 24-29 hours ahead into a continuous day-ahead forecast time series. We use the System Advisor Model (<https://sam.nrel.gov/>) to convert solar irradiance to power.

Our initial plan was to use NAM to develop day-ahead forecasts for each balancing region in the study area and use data from the National Solar Radiation Database (NSRDB) for “actual” 15-minute (real-time) solar profiles (interpolated from 30-minute data) and for calculating persistence forecasts. However, differences between NAM day-ahead forecasts and NSRDB 15-minute solar data were large, even after de-biasing the NAM data, leading to forecast errors for the NAM forecast that were significantly higher than the persistence forecast. To obtain better consistency between day-ahead forecasts and real-time data, we chose to use NAM to generate both the day-ahead forecasts and 15-minute solar profiles. We then calculated persistence forecasts using the NAM 15-minute solar profiles.

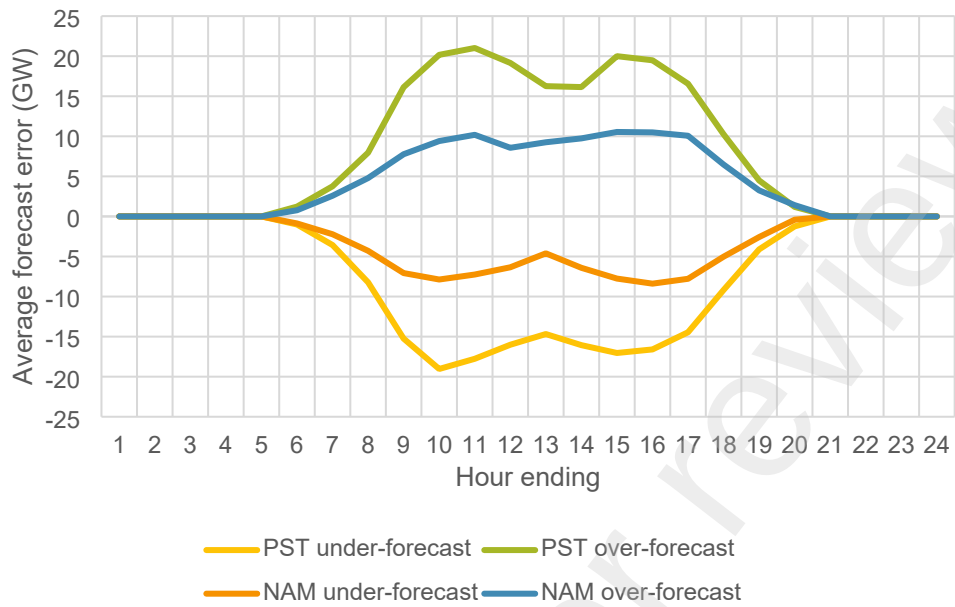
A.3. Forecast Errors

The day-ahead solar forecast errors in this study have several desired characteristics. (1) Forecast errors are lower for the total region (2.8% NAM / 5.8% PST NMAE) than for individual balancing regions (3.7%-4.6% NAM / 7.9%-11.1% PST NMAE) (Table A.3.1). (2) The difference between NAM and persistence forecast errors (forecast improvements) is not uniform across the daily solar cycle (Figure A.3.1). (3) Forecast errors and forecast improvements are complex. Examples of forecast complexity are that NAM forecast errors do not have the same sign as persistence errors (26% of hours), that in some hours the persistence forecast performs better than the NAM forecast (17% of hours), and that many days have large hourly under-forecasts and over-forecasts that occur on the same day. For the last of these, in 23% of days with the NAM forecast and in 12% of days with the persistence forecast the daily sums of hourly under-forecast errors and over-forecast errors are within 50% of each other, indicating that the daily over- and under-forecasts are close in magnitude.

Table A.3.1. Solar Forecast Error Metrics for Balancing Regions

	DUK	LGEE	SOCO	TVA	Total
NAM forecast					
NMAE	3.7%	4.6%	4.6%	4.1%	2.8%
nRMSE	7.9%	11.1%	10.5%	9.4%	5.8%
Persistence forecast					
NMAE	7.0%	8.3%	7.3%	7.5%	5.5%
nRMSE	13.9%	17.4%	15.5%	15.6%	10.8%

568



569

570 **Figure A.3.1. Average daily NAM and persistence forecast errors (under-forecast and over-forecast).**

571 Table A.3.2 shows day-ahead (24-hour-ahead) and day-ahead lookahead (48 and 72 hours ahead of the
572 forecast interval) solar forecast errors for the persistence and NAM forecasts, with both hourly and 4-
573 hour block (sum of 4 hours) forecast errors for the longer-term lookahead forecasts. The longer-term
574 forecast errors are significantly higher than the 24-hour-ahead forecast errors, as expected. The 4-hour
575 block errors, which are what the production simulation model optimizes for, are significantly higher than
576 hourly errors for the longer-term forecasts. The NAM forecast provides a nearly 50% improvement in
577 forecast accuracy for the 24-hour-ahead forecast but only a 40% improvement for the longer-term
578 forecasts.

579 **Table A.3.2. 24, 48, and 72-Hour-Ahead Solar Forecast Errors (NMAE)**

Forecast	Time resolution	Persistence forecast	NAM forecast
24-hour-ahead	Hourly	5.5% (9.7%)	2.8% (4.9%)
48-hour-ahead	Hourly	10.1% (16.6%)	6.0% (10.6%)
	4-hour block	38.3% (55.8%)	22.8% (33.3%)
72-hour-ahead	Hourly	6.1% (10.8%)	10.3% (16.9%)
	4-hour block	38.9% (56.7%)	23.2% (33.9%)

580 Notes: The NMAE metrics in this table include periods with zero values (values excluding zero values are in
581 parentheses).

582 **Appendix B. Additional Results**

583 **B.1. Cost Driver Regression**

584 To explore drivers of changes in costs, we regressed the change in daily costs between NAM and PST
585 (NAM – PST) against daily solar availability (PV_i for day i), daily load (LD_i), daily PST forecast error (FE_i),

daily forecast improvement (FI_i), and dummy variables for spring (SP_i), summer (SU_i), fall (FA_i), and winter (WI_i).

$$DC_i = \alpha + \beta_1 PV_i + \beta_2 LD_i + \beta_3 FE_i + \beta_4 FI_i + \beta_5 SP_i + \beta_6 SU_i + \beta_7 FA_i + \beta_8 WI_i + \varepsilon \quad \text{Eq. (B.1)}$$

Table B.1.1 shows coefficients and p-values from the regression. The adjusted R^2 value was 0.29, suggesting that these variables do not explain a large amount of the variance in daily changes in costs between the NAM and PST cases.

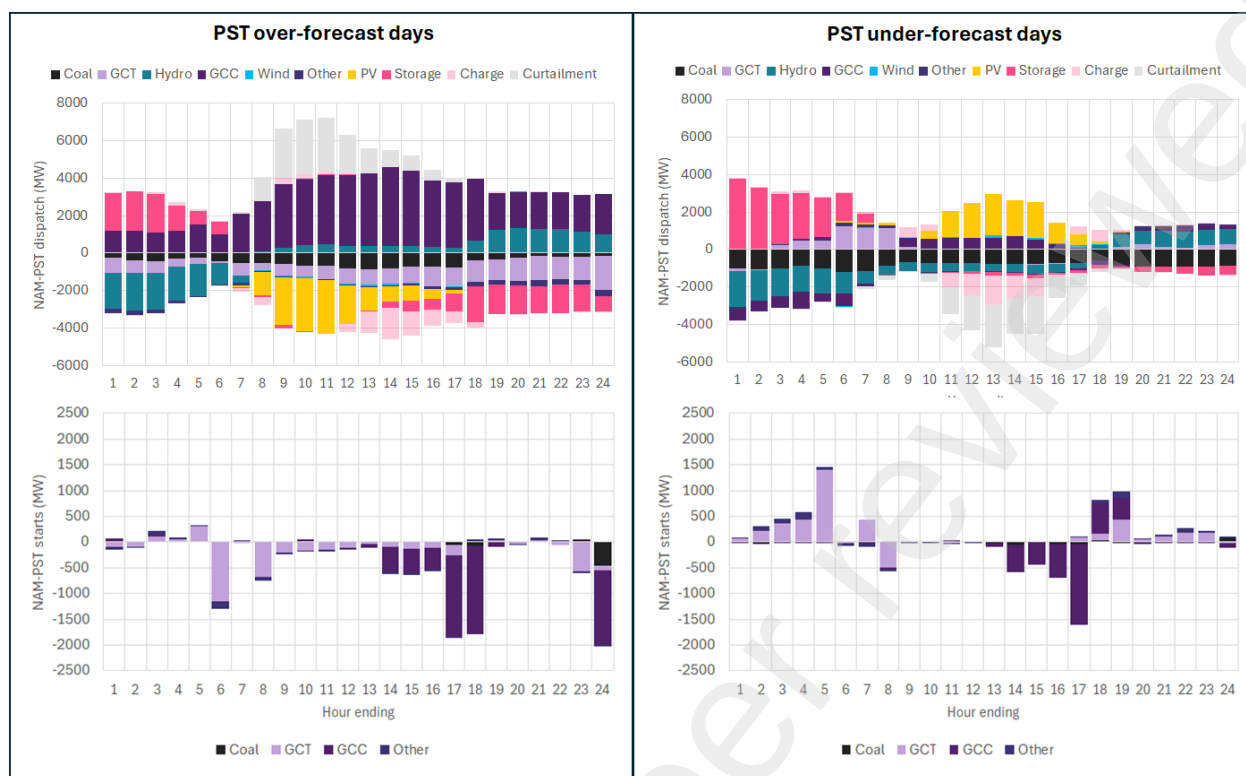
Table B.1.1. Regression Results

Variable	Coefficient	P-value
Daily solar availability (PV)	0.0014	0.00000
Daily load (LD)	-0.0010	0.00006
Daily PST forecast error (FE)	-0.0012	0.00000
Daily forecast improvement (FI)	-0.0020	0.00000
Spring (SP)	-0.3313	0.01278
Summer (SU)	-0.4463	0.00931
Fall (FA)	-0.1138	0.35044
Winter (WI)	-0.0288	0.84709

B.2. Unit Commitment and Dispatch for the HighCoord and FlexGas Scenarios

HighCoord Scenario

Figure B.2.1 shows changes in real-time average daily dispatch (top panels) and day-ahead and real-time average daily starts (total MW, bottom panels) between the NAM and PST cases in the HighCoord scenario. On days with a PST over-forecast (actual solar < forecast, lefthand side of Figure B.2.1), the results are structurally similar to but larger than those in the Base scenario. On days with a PST under-forecast, the main difference between the NAM and PST cases is in solar generation, storage charge and discharge, and hydro generation. Differences between the HighCoord and Base scenarios are due to lower hurdle rates, which create differences between the NAM and PST cases both in day-ahead schedules and real-time dispatch. In the HighCoord scenario, both the NAM and PST cases have 6% more solar generation than in the Base scenario, which is likely an important driver of differences between the two scenarios.



Notes: A positive (negative) value for generation, storage discharge, and curtailment means that they are higher (lower) in the NAM than in the PST case. A positive (negative) value for storage charging means that it is lower (higher) in the NAM than in the PST case.

Figure B.2.1. Changes in average daily dispatch (top plots) and starts (bottom plots) between the NAM and PST cases for PST over-forecast and under-forecast days, HighCoord scenario

Table B.2.1 shows changes in annual real-time generation, storage charging, and curtailment between NAM and PST for the HighCoord scenario. Changes in coal, GCC, and GCT generation are larger than in the Base scenario. The large increase in GCC generation in the NAM case (+12.3 TWh) corresponds to lower coal (-5.7 TWh), GCT (-3.1 TWh), solar (-1.2 TWh), and oil and gas steam (-0.5 TWh) generation relative to the PST case. The NAM case has higher storage discharge (+1.4 TWh) but less solar generation, implying that higher storage discharge is being charged mainly by GCC generation. The NAM case has higher curtailment than the PST case (+1.2 TWh).

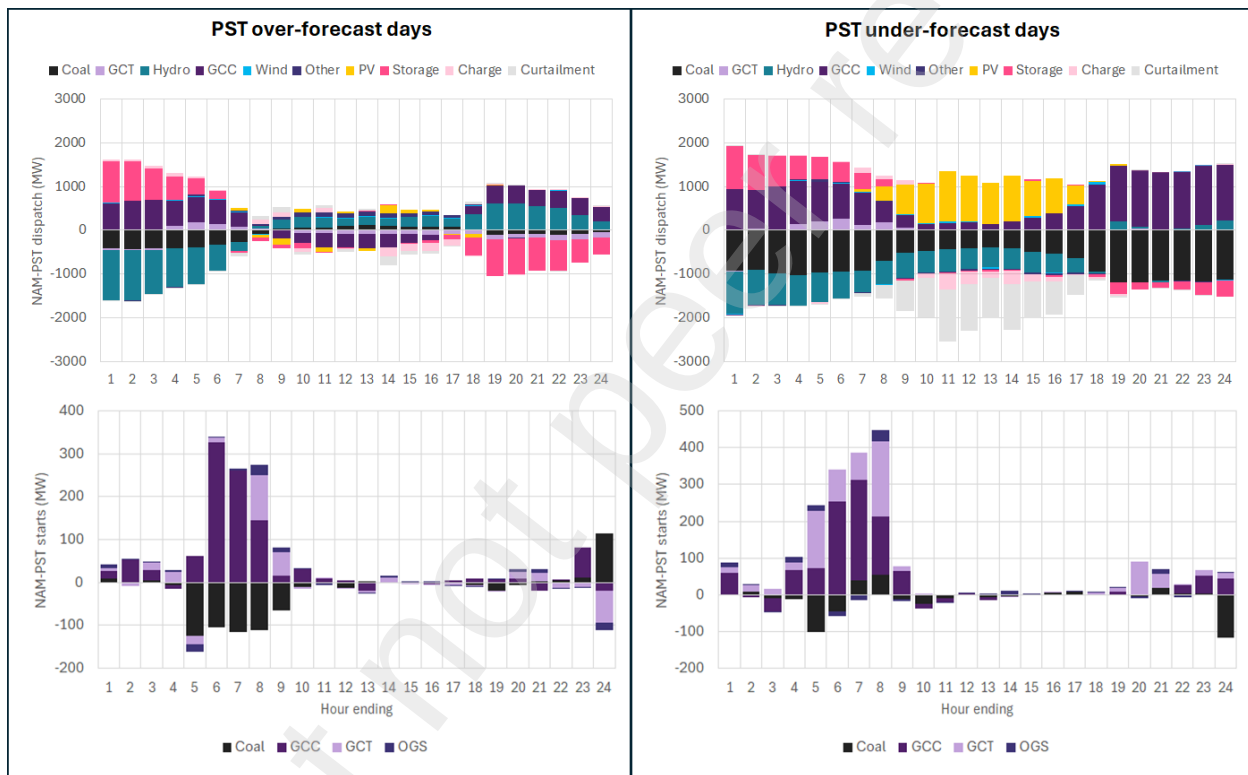
Table B.2.1. Annual Generation (TWh/yr) in the NAM and PST Cases, Change in Annual Generation (Δ , NAM-PST), and Percentage Change ($\%\Delta$) in Annual Generation, HighCoord Scenario

	Coal	GCC	GCT	OTH	NUC	HYD	WND	PV	STD	STC	CURT
NAM	46	120	8	6	216	31	22	250	72	77	55
PST	52	108	11	5	216	33	22	251	70	76	54
Δ	-5.7	12.3	-3.1	-1	0.0	-1.4	0.0	-1.2	1.4	1.5	1.2
$\%\Delta$	-11%	11%	-28%	-10%	0%	-4%	0%	0%	2%	2%	2%

Notes: OTH is other thermal, including oil and gas steam and biomass; NUC is nuclear; HYD is hydropower, WND is onshore and offshore wind; PV includes distributed PV; STD and STC are storage discharge and charge, respectively; CURT is curtailment.

FlexGas Scenario

Figure B.2.2 shows changes in real-time average daily dispatch (top panels) and day-ahead and real-time average daily starts (total MW, bottom panels) between the NAM and PST cases in the FlexGas scenario. Changes in generation and starts are significantly smaller in the FlexGas scenario than in the Base scenario, which implies that enabling more flexibility in GCC commitment narrows the differences between the NAM and PST cases. On days with a PST over-forecast, differences in coal and GCC commitment drive differences in storage dispatch and hydro generation. On under-forecast days, in the NAM case the model absorbs more solar generation, discharges more storage in the mornings, has less hydro generation, and commits and dispatches GCCs rather than coal generation. The changes in coal commitment in the FlexGas scenario are identical to those in the Base scenario. The main difference in commitment between scenarios is with GCC units.



Notes: A positive (negative) value for generation, storage discharge, and curtailment means that they are higher (lower) in NAM than PST. A positive (negative) value for storage charging means that it is lower (higher) in NAM than in PST.

Figure B.2.2. Changes in average daily dispatch (top plots) and starts (bottom plots) between the NAM and PST cases for PST over-forecast and under-forecast days, FlexGas scenario

Table B.2.2 shows changes in annual real-time generation, storage charging, and curtailment between the NAM and PST cases for the FlexGas scenario. Changes in generation, storage charging, and curtailment are generally on par with those in the Base scenario, with slightly higher GCC generation, less coal generation, and less storage discharge in the NAM case. The main contributor to differences in production costs between the FlexGas and Base scenarios is differences in GCC and coal starts and start costs rather than differences in annual generation. The reason why the NAM case has slightly higher

costs than the PST case in this FlexGas scenario likely has to do with sub-optimal coal and gas commitment in the NAM case. In some situations, the model may de-commit coal and commit gas units, based on the NAM forecast and its reserve requirements, when the cost optimal solution, per the PST forecast and its reserve requirements, would have been to keep coal units online. Coal commitment in the NAM case is identical in the FlexGas and Base scenarios. Because differences in between the NAM and persistence forecasts are not uniform across time and commitment decisions incorporate multi-day forecasts, it is difficult to isolate when and why sub-optimal commitment may be occurring. Extending the lookahead period in the FlexGas scenario did not change this result. Solar forecasting and coal unit commitment is an important area for future study.

Table B.2.2. Annual Generation (TWh/yr) in the NAM and PST Cases, Change in Annual Generation (Δ , NAM-PST), and Percentage Change ($\%\Delta$) in Annual Generation, FlexGas Scenario

	Coal	GCC	GCT	OTH	NUC	HYD	WND	PV	STD	STC	CURT
NAM	61	107	7	4	216	32	23	249	72	77	68
PST	65	103	7	4	216	33	23	248	72	76	69
Δ	-3.7	3.8	0.0	0.1	0.0	-1.6	0.1	1.4	0.3	1.5	-1.1
$\%\Delta$	-6%	4%	0%	3%	0%	-5%	0%	1%	0%	2%	-2%

Notes: OTH is other thermal, including oil and gas steam and biomass; NUC is nuclear; HYD is hydropower, WND is onshore and offshore wind; PV includes distributed PV; STD and STC are storage discharge and charge, respectively; CURT is curtailment.

As a one-off sensitivity, we ran the FlexGas scenario with an extended day-ahead lookahead period (48 hours beyond the day-ahead interval), to observe whether the NAM case would still have higher production costs than the PST case. Cost savings in the PST case (relative to the NAM case) increased in this scenario (from \$41 million in the FlexGas scenario to \$54 million in this sensitivity), suggesting that extending the day-ahead lookahead period may not address the issues that are driving higher NAM case production costs in the FlexGas scenario.

B.3. Real-time Reserve Violations

Table B.3.1 shows real-time reserve violations for the Base scenario. The frequency and magnitude of real-time reserve violations are small and are similar between the NAM and PST cases.

Table B.3.1. Real-Time Reserve Violations in the NAM and PST Cases, Base Scenario

	Contingency spinning	Contingency non-spinning	Regulation
NAM			
Frequency (number of intervals, % intervals)	50 (0.1%)	21 (0.1%)	94 (0.3%)
Total violations (GW-h/yr)	4	8	6
Max violation (GW)	2	5	2
PST			
Frequency (number of hours, % annual hours)	59 (0.2%)	12 (0.0%)	91 (0.3%)
Total violations (GW-h/yr)	3	3	4
Max violation (GW)	2	5	1

B.4. Extended Real-Time Lookahead

The use of a 24-hour rather than a 12-hour lookahead period in the longer real-time lookahead (LongRTLook) sensitivity results in significantly higher production cost savings (\$152 million per year in savings) relative to the Base scenario (\$87 million per year in savings) (Table B.4.1). The longer real-time lookahead reduces costs in both the NAM and PST cases. In both cases, higher solar generation (less curtailment, +10 TWh/yr solar generation in NAM and +8 TWh/yr in PST) displaces mostly GCC and GCT generation (-5 TWh/yr in NAM, -6 TWh/yr in PST). Incremental solar generation is higher in the NAM case, which explains the additional cost savings from higher forecast accuracy (NAM minus PST) with the extended real-time lookahead.

B.4.1. Annual Production Costs (Million 2023\$) and CO₂ Emissions (MtCO₂) in the Base Scenario and LongRTLook Sensitivity

	BaseLook		LongRTLook	
	NAM	PST	NAM	PST
Start costs (M\$)	\$1,223	\$1,323	\$1,025	\$1,133
Fuel costs (M\$)	\$6,999	\$6,976	\$6,719	\$6,746
Variable O&M costs (M\$)	\$901	\$911	\$868	\$884
Total costs (M\$)	\$9,124	\$9,211	\$8,612	\$8,764
CO ₂ emissions (MtCO ₂)	112.5	114.0	108.2	110.3

Our real-time lookaheads use perfect foresight (actual rather than forecasted values). The results in Table B.4.1 suggest that extending the real-time lookahead period may increase the cost savings from day-ahead solar forecast improvements, but they do not shed light on whether an improved solar forecast would enhance the benefits of a longer real-time lookahead.