NY-Sun Solar Photovoltaic Program Performance Persistence Study

Final Report

Prepared for:

New York State Energy Research and Development Authority
Albany, NY

Dana Nilsson
Senior Project Manager, NYSERDA

Prepared by:

DNV

Corporate Headquarters: Katy, TX

Miriam Goldberg, Sr. Principal, DNV

Bradley Campbell, Principal, DNV

Kora Dreffs, Sr. Consultant, DNV

Dale Tutaj, Sr. Engineer, DNV

Antonia Sanhueza, Professional Consultant, DNV

NYSERDA Contract 177993 June 2025

Record of revision

Document Title
NY-Sun Solar Photovoltaic Program Performance Persistence June 2025

Revision Date	Description of Changes	Revision on Page(s)
June 2025	Original Issue	Original Issue

Notice

Table of contents

List of figures

List of tables

Glossary of key terms and acronyms

Annual production. The amount of electricity generated by a PV system in a year, typically measured in kWh.

Normalized annual production. The amount of electricity a PV system would have generated in a given year, if the irradiation levels (cloud cover) over the year had been typical. Normalized annual production is determined from the system's actual production in each month and the normal-year irradiance for its location. This correction does not account for snowfall covering the panels or changes in shading of nearby structures that might affect solar energy production.

Capacity factor (CF). The ratio of a system's annual production to the production that would result if the system could produce at its full capacity for every hour of the year. For this report, all capacity factors are calculated using normalized annual production.

Performance loss rate (PLR). The year-over-year percentage reduction in normalized annual production. The PLR estimate assumes the same underlying proportional reduction in CF over successive years for the period over which the PLR is calculated. The PLR is calculated using actual production and does not account for the expected or modeled production.

Persistence. The extent to which first-year impacts determined for a program continue to exist in later years. In this study, persistence is assessed in terms of the system performance loss rate for operating communicating systems.

Executive summary

The NYSERDA NY-Sun PV Incentive Program provides cash incentives and/or financing according to a megawatt block structure, with a goal of installing 10 GW of PV capacity by 2030. Solar installations typically remain in place for 25 years or more. Planning the future grid requires knowing both total installed solar capacity and how well system production holds up over time. This study analyzes historical production data for systems installed in New York State (NYS) to estimate the performance loss rates.

This study estimates an average annual **performance loss rate** (PLR) of 0.83% ± 0.09% across New York State. Aggregated over time and New York's solar fleet, an annual PLR of 0.83% would mean a reduction of more than 20% in the annual production by 2050 from the 10 GW of PV capacity targeted under New York's Climate Act Scoping Plan. This represents a loss of around 2,200 GWh, underscoring the importance of incorporating degradation into long-term energy planning. Even considering this PLR, New York could still achieve approximately 80% of the expected annual energy production by 2050 from the 10 GW of solar capacity planned for installation by 2030 under the Climate Act Scoping Plan.

The analysis reveals significant regional and system-level variation. PLRs range from 0.5% per year in Long Island to 1.5% per year in Upstate New York. Purchased systems and Power Purchase Agreement (PPA) systems exhibit higher loss rates than leased systems, monocrystalline modules outperform polycrystalline, and systems using microinverters show lower PLR than string and optimizers. Finally, systems in Disadvantaged Communities (DACs) show higher loss rates, potentially due to more flat roofs.

Analysis results will be used to help the state track progress on its decarbonization goals and identify opportunities to mitigate performance risks across the solar portfolio.

Approach

The study analyzes monthly production data for a sample of NY-Sun projects that were completed between February 2012 through November 2021. Some of the projects have been included in NY-Sun first year impact studies. This was supplemented with similar data for a set of NY residential solar PV projects that the evaluation contractor has reviewed for other clients.

Based on the available data, the study estimates that a large fraction of these supplemental projects, possibly the majority, received NY-Sun incentives.

Methods similar to those used for the impact studies combined each project's monthly production data with same-period and normal-year radiance data to produce irradiance-normalized production, the production that would have been seen for that project and month in a year with typical cloud cover and related conditions. The monthly normalized production was expressed as a capacity factor (CF), that is, as a fraction of the production that would occur if the project could produce at full capacity for all hours of the month.

The PLR was calculated as the year-over-year reduction in capacity factor, controlling for month of the year. The PLR was calculated for individual projects and by project subgroup. The subgroups analyzed were region (Upstate, Downstate, and Long Island), sector (residential and nonresidential), size (<200 kW, 200 to 750 kW, and >750kW), DAC category, purchase type (lease, PPA, purchase, and unknown), module family (monocrystalline, polycrystalline and unknown), inverter type (microinverter, optimizer, sting, and unknown), inverter manufacture, and panel manufacturer.

The primary analysis in this study assesses the PLR for operating, communicating projects. "Operating" refers to projects with recorded production values and "communicating" are projects that report the production. This analysis excluded months with missing or zero production. The available data does not distinguish between periods of disrupted production versus periods of continuing production with disrupted communication to the project owner. Directly analyzing performance loss including months with data could therefore overstate the level of production decline. A supplemental analysis indicated that, within the study timeframe, the incidence of no- and low-production months is in fact quite low; thus, it does not appear that excluding these months results in appreciable understatement of the effective loss rate including temporary non-performance.

The key analysis elements are:

- **Pooled PLR by subgroup and month**, calculated across all projects within a subgroup, separately for each month of the year. The main PLR estimates of the study shown in Table ES-1 use this method, weighting the monthly PLR values by the monthly CFs to produce an overall weighted result.

- **Project-specific PLR**, calculated for each project across all months and years of production data for that project. The average project-specific PLR by subgroup was the primary basis for comparing PLR of different subgroups.
- **Cross-factor analysis** to identify the effect of particular project features, with all other features held constant. This analysis was used to mitigate the effects of different mixes of characteristics within subgroups when comparing PLR between subgroups.

The pooled PLR estimates and cross-factor analysis generally corroborated the indications from the simpler comparisons of project-specific averages, with the cross-factor analysis providing additional insights in key cases.

Results

Performance loss rates by region

Table ES-1 shows the PLR by region, as well as overall. The results are well determined, with standard errors between 9% and 32% of the estimate. The overall PLR of 0.83% per year is statistically significantly different from the benchmark of 0.64% at the 10% significance level. The table shows a much higher PLR for the Upstate region, and a much lower value for Long Island. Reasons for these differences are not clear, but may be related to more severe weather conditions Upstate leading to a higher level of equipment degradation.

Table ES-1. PLR by region and overall
CF-weighted pooled PLR by region and month

Region	Number of projects	Avg capacity (kW)	Avg Capacity Factor	PLR	Standard error of PLR
Downstate	1,670	22	12.9%	0.87%	0.19%
Long Island	4,946	10	13.2%	0.41%	0.13%
Upstate	2,365	81	12.0%	1.67%	0.15%
All	8,981	31	12.7%	0.83%	0.09%

PLR comparisons across other subgroups

For most comparisons across subgroups, the average of the project-specific PLRs in that subgroup is used for the comparative discussions, as a more intuitive basis for comparison. However, the pooled monthly estimates weighted by CF are more meaningful at the regional level. The project-specific averages are used in most of the discussion that follows of factors associated with higher and lower PLR, with attention to the potential effects on such comparisons of the overstated Upstate PLR. The cross-factor PLR is used to further evaluate the observed trends.

PLR by DAC and non-DAC location

A direct comparison of average project-level PLR found a significantly lower PLR in DAC than in non-DAC geographies. However, this difference appears to be related to the regions where the DAC and non-DAC projects fall. In the study pool, 82% of the DAC projects are in Long Island, where the PLR is low, and only 7% are Upstate, where the PLR is high. Non-DAC projects, by contrast, are 42% in Long Island and 36% Upstate. For the cross-factor comparison, DAC has a 0.48% higher PLR then non-DAC projects, when controlling for regional and other differences in characteristics.

PLR by purchase type

A total of 82% of projects are leased. Power purchase agreement (PPA) systems (15% of projects in study) exhibit much higher PLR than leased ones. Purchased systems are relatively few in the study data set (1%) and have much higher PLR than leased. The cross-factor analysis confirmed these directional comparisons with PPA having a PLR 2.28% higher than leased projects and purchased projects having a 3.36% higher PLR, when controlling for other factors. While the study cases are heavily weighted towards leased projects, since 2012 the NY-Sun Program has had about 27% leased, 10% PPA and 63% purchased systems.

Purchased systems are likely to have less consistent maintenance, since owners may not realize the systems are not performing well, and may not call for repairs immediately. Thus, the higher PLR for these systems makes sense.

PLR by module family

Monocrystalline systems exhibit very low PLR, not statistically different from 0. By contrast, polycrystalline systems exhibit a PLR somewhat higher than the overall average. The polycrystalline systems are generally older, and there may be reasons other than the module type that they have higher performance loss rates. Over one-third of the systems have an unknown module type. These have a PLR similar to the polycrystalline systems. The very low PLR for monocrystalline systems observed in the project-level averages was borne out by the cross-factor analysis. The cross-factor analysis shows that polycrystalline modules have a 1.20% higher PLR compared to monocrystalline, while unknown module types are 1.44% higher compared to monocrystalline.

PLR by inverter type

Microinverter systems appear to have higher performance loss rate than those with optimizers or string inverters, even controlling for other characteristics. Microinverters have a 0.96% higher PLR than Optimizer systems and a 2.5% higher PLR than string inverters. A possible explanation is that an individual microinverter failure within a system results in only moderate performance decline and may go undetected and unaddressed; whereas a string failure would bring production to zero, hence be more likely to be promptly repaired.

Findings and recommendations

Finding 1

Performance loss rates estimated in this study represent the performance loss of operating and communicating systems, over roughly the first four to seven years of life. Temporary outages of systems or communications in principle could contribute to additional production loss across the NYS solar fleet. The frequency of such disruptions across the data and time span studied is quite low, and does not suggest the need to consider temporary disruptions as an additional loss factor over this time horizon.

Recommendation 1:

Consider a study to investigate the proportions of incentivized systems that are still in operation at different ages while considering technical features (e.g., whether they are monocrystalline or polycrystalline or tracked vs. fixed systems). Such a study could include working with the system managers to identify rates of data loss and reasons.

NYSERDA Response to Recommendation: Pending. NYSERDA will consider this study as part of future performance analyses.

Finding 2

Overall performance loss rates are estimated at 0.83% per year across projects in New York State. This rate is statistically significantly higher than the benchmark of 0.64% generally used by the contractor for project pool assessments on behalf of developers. If the current loss rate were to persist through the system lifetimes, New York could still achieve approximately 80% of the expected annual energy production by 2050 from the 10 GW of solar capacity planned for installation by 2030 under the Climate Act Scoping Plan.

The estimated loss rate varies substantially across regions, with lower loss rates in Long Island and higher rates in Upstate New York. These regional differences may be related to differences in weather conditions that may cause damage to equipment or differences in snow accumulations.

Recommendation 2:

Take estimated performance loss rates by region into account when projecting future production of NYS solar systems as a whole.

NYSERDA Response to Recommendation: Rejected. NYSERDA will consider the estimated performance loss rates when projecting future production of NYS solar systems as a whole.

Finding 3

Both purchased systems and, to a lesser extent, PPA systems have higher PLR than leased systems. This result is expected for purchased systems, where maintenance may be worse due to less active monitoring. PPA and leased systems are typically maintained by third parties who are responsible for the performance of the systems.

Recommendation 3:

Consider offerings that could support improved maintenance for purchased systems.

NYSERDA Response to Recommendation: Pending based on continued NYSERDA program funding

Finding 4

Monocrystalline modules are associated with lower performance loss rates than polycrystalline. These empirical findings are consistent with expectations for these technologies. Newer systems are more likely to be monocrystalline, which have a higher persistence in production over the life of the system.

Recommendation:

No recommendation, given that most new installations will be monocrystalline.

NYSERDA Response to Recommendation: N/A

Finding 5

Microinverter systems have higher performance loss rates than those with optimizers or string systems, even controlling for other characteristics. A possible explanation is that failures of individual microinverters within a system go undetected and unaddressed resulting in higher performance loss rates. In contrast, a string system failure would result in zero production and could be more likely to be recognized and repaired.

Recommendation 5:

Consider offerings that would support system owners identifying and addressing reduced performance that may be related to microinverter failures. One option could be to encourage continued sharing of production data with NYSERDA, in exchange for NYSERDA sending alerts for indications of maintenance needs. Such an approach could provide NYSERDA with an expanded data series for future persistence studies.

NYSERDA Response to Recommendation: Rejected. NYSERDA does not have the ability to implement this recommendation.

Finding 6

Controlling for other characteristics, systems in DAC locations have higher PLR than those in non-DAC areas.

Recommendation 6:

Consider additional research to determine factors associated with greater in DAC areas, for both third-party operators and homeowners with purchased systems.

NYSERDA Response to Recommendation: Pending. NYSERDA will consider completing additional research to determine which factors are most closely associated with a greater PLR in DAC areas.

Finding 7

Most projects analyzed in this study had five or fewer years of production data. Only a handful had more than seven years. As a result, the performance loss rates developed in this study may not be indicative of loss rates after 10 or more years.

Recommendation 7: Continue to collect production data for the NY-Sun sample included in this study, to establish a larger data set of longer production records that can be used for further study. Explore processes that could be used to collect later production data for the supplemental data sets included.

NYSERDA Response to Recommendation: Pending. NYSERDA will consider this study as part of future performance analyses and/or by studies completed by reputable third parties.

1 Introduction

This report presents a persistence study of solar photovoltaic (PV) projects installed in New York State (NYS) from February 2012 through November 2021. A substantial fraction of these projects were supported by NYSERDA's NY-Sun program. Prior evaluations and studies of that program conducted by the evaluation contractor team[1] examined program impact and first-year annual capacity factors. This study examines the persistence of installed system production performance and its relationship to the characteristics of the systems.

1.1 NY-Sun Program description

The NYSERDA NY-Sun PV Incentive Program, open August 12, 2010, through December 29, 2025, provides cash incentives and/or financing according to a megawatt (MW) block structure. "Blocks," or specific MW targets per defined sector and geographic region of New York, are active on a rolling basis until fulfilled. The program expanded its original goal of installing 3 gigawatts (GW DC) of PV capacity by 2023 to 6 GW DC by 2025, and NYSERDA's 2019 petition to extend the NY-Sun program and increase funding was approved in 2020. On September 20, 2021, Governor Kathy Hochul called for the expansion of the NY-Sun Program, with a goal of 10 gigawatts by 2030. NYSERDA and DPS developed the Solar Roadmap, filed December 17, 2021. An April 2022 order adopted the roadmap recommendations. The order expanded funding for base incentives, the Solar Energy Equity Framework, and incentive adders.

1.2 Existing studies and assumptions

A white paper by DNV (Appendix A[2]) has reviewed available studies of solar system performance over time. The studies reviewed include various technologies, numbers of systems, system sizes, regions, and vintages. The system-level degradation determined from these studies ranges from 0.2% to 1.2% per year, with an average of 0.62%. For its assessment practice, DNV uses a value of 0.64% per year as its standard, based on a key study among those reviewed in the paper, with exceptions for certain specific technologies. The present study treats 0.64 as a useful

---

[1] DNV for NYSERDA, *NY-Sun Solar Photovoltaic Program Impact Evaluation for Systems Installed April 1, 2018 through March 31, 2021*, June 2024. https://www.nyserda.ny.gov/-/media/Project/Nyserda/Files/Publications/PPSER/Program-Evaluation/NYSERDASolarPhotovoltaicImpactEvaluationReport.pdf

[2] Henry Hieslmair, *DNV's Views on Long-Term Degradation of PV Systems*, May 2024. https://www.dnv.com/publications/dnv-views-on-long-term-degradation-of-pv-systems/

benchmark, while recognizing that different groups of projects may have different overall degradation rates, and individual systems vary even more.

NY-Sun to this point has not focused on projecting long-term production levels and has not adopted a performance loss rate assumption. New York Green Bank, as a financing entity, does use an assumed loss rate, in its risk analyses, currently 0.5% to 0.75%.

An annual performance loss rate of less than 1% per year is small compared to the variability of actual production year to year. However, aggregated over New York's PV fleet and over time, a sustained performance loss rate of 0.64% per year would mean a nearly 15% reduction in annual production over a life span of 25 years. For the 10 GW of PV capacity targeted to be installed by 2030 under New York's Climate Act Scoping Plan, it would mean a loss on the order of 1,700 GWh of annual production by 2050.

The present study is the first to focus specifically on projects in New York State, with New York's climate and PV system characteristics. This study uses empirical data from a large pool of systems installed in the state over many years.

1.3 Evaluation objectives and methods

1.3.1 Objective and key study elements

Table 1-1 indicates the objective of this study. Table 1-2 indicates the key parameters studied to meet that objective.

Table 1-1. Study objective, purpose, and method

Objective	Purpose	Method
Assess the persistence of solar production over time, and how it varies by solar system characteristics	Allow NYS to estimate the contributions of previously installed customer-sited projects to current and future renewable energy generation in the state	Compile production data from multiple years from a large pool of NYS solar PV projects, and analyze the change in production over time

Table 1-2. Parameters studied

Parameter	Uses	Approach
Supplied power (nameplate kW DC)	Provide power supplied per site and region.	Collect nameplate DC capacity (kW) for sites from tracking data
Energy impact (kWh annual production),	Assess irradiance-normalized production and change in production over time	Collect production data and calculate normalized production under typical weather conditions for each year of production, for each project.
Capacity factor (%)	Normalizes production data by expressing production for a given time period as a percent of hypothetical production if the system operated at full nameplate capacity continuously over the same period of time.	Calculate site-level capacity factors based on available nameplate and irradiance-normalized production data for each year.
Change in performance over time	Identify how much of initial production is retained over time	Analyze the change in irradiance-normalized capacity factor over time

Electricity production from a PV system in a given year depends in part on the weather, in particular the levels of snow and cloud cover, over the year. To remove some of the variability from the analysis of changes over time, each project's observed production data for each year is used to calculate the production that project would have had in a year of normal irradiance. This calculated normal-year production is the irradiance-normalized production for the project and year. The normalization does not adjust for snowfall levels.

Persistence in this study is assessed in terms of the *performance loss rate* (PLR). The PLR is the year-over-year percentage reduction in normalized annual production.

1.3.2 Data sources

The evaluation contractor team obtained production data from New York's Distributed Energy Resources (DER) portal, and from system installers and project managers for a sample of NY-Sun projects. Data was requested for a sample rather than comprehensively to limit the burden on the installers and managers. To develop multi-year histories for this study, data from the sample was collected periodically from the providers. Additional project meta-data was collected for the sample from NY-Sun Program tracking data and Open NY Solar. All production data available in the DER portal for installations through 2017 was included in the study. For many projects, this covers only the first three or four years of production, after which reporting to DER stops.

Over time, some installers and project managers who had provided production data went out of business, or for other reasons stopped responding to requests for data. As a result, long series are not consistently available for the earliest sampled sites. Later installations necessarily have shorter data series to this point. The NY-Sun team supported the evaluation contractor team's efforts to obtain as much data as possible from their sampled projects.

To provide a richer data set for this study, the data from this sample of NY-Sun/DER projects was supplemented by a data set of multi-year production data from residential projects in NYS that the evaluation contractor had aggregated in the course of other work, supporting project owners for purposes of refinancing. The supplemental project-level production data was made available for use in this study, with the stipulation that only aggregate results developed using those data may be made available to NYSERDA or the public. The supplemental data varies in manufacturers, technologies, and other system characteristics, which helps ensure broader applicability. For a detailed comparison of characteristics between NY-Sun sample and NYS supplemental data, please refer to Appendix B.

Projects in the supplemental data set were matched against the NY-Sun/DER impact sample to ensure there were no duplicate projects in the combined data set. No duplicate cases were identified, as described further in Section 2.1, Study data.

Table 1-3 indicates the production data used for the study.

Table 1-3. Data sources used in the study

Source	Sector	Size category	Project capacity (kW)	Number of projects	Time frame of installations	Time frame of production data	Median months of production data
NY DER Portal	Non-residential	<200	55–200	30	2012–2017	2012–2024	69
NY DER Portal	Non-residential	200–750	200–743	133	2012–2018	2012–2024	83
NY DER Portal	Non-residential	≥750	783–2,994	79	2013–2018	2013–2024	83
NY-Sun installation contractors	Non-residential	<200	6–189	33	2012–2016	2013–2024	59
NY-Sun installation contractors	Non-residential	200–750	200–348	6	2016	2016–2022	34

Source	Sector	Size category	Project capacity (kW)	Number of projects	Time frame of installations	Time frame of production data	Median months of production data
NY-Sun installation contractors	Residential	<200	3–17	105	2012–2017	2013–2024	88
DNV proprietary NYS data	Residential	<200	1–37	13,587	2012–2023	2011–2023	46

Table 1-3 indicates that projects used in this study are up to 12 years old as of this analysis. The number of years of production data may be less than the project age, as noted. The analysis used here assesses the average rate of change, in proportional terms, across all available time periods. With this approach, all year-over-year changes for an individual site contribute equally to the analysis. This allows the use in the analysis of projects with long or short production data series, making the fullest use possible of the available data.

1.3.3 Solar persistence as addressed in this study

The change in output from an installed solar system over time can result from various causes, including factors related to system connectivity and dispatch, panel shading and temporary soiling, and degradation of system components. A detailed discussion of these factors is included in the white paper provided in Appendix A of this report.

The present study assesses the performance loss rate in terms of the rate of reduction in irradiance-normalized production from functioning, communicating systems. This performance loss rate is essentially a system degradation rate, as described in the white paper.

The white paper reviews findings from many studies, addressing systems installed from 2008 through 2020 across the United States and in other countries. That review provides a robust basis for considering performance loss rates. As noted, the present study uses the white paper’s suggested rate of 0.64% per year as a reference point for the findings here. The prior work referenced in the white paper is not specific to New York State. The studies referenced are mostly based on field data but include some lab studies. Nonetheless, this operational assumption for ongoing assessments serves as a useful benchmark of PLR.

The present study assesses performance loss rates specific to systems installed in New York State. A large fraction of these systems were installed with incentives from NY-Sun.

2 Results, findings, and recommendations

2.1 Study data

2.1.1 Overlap of data sources

To identify potential duplicate cases between NY-Sun/DER impact sample and the supplemental data, the evaluation contractor team attempted to match the more limited NY-Sun residential impact sample projects to projects in the supplemental data, based on location and other characteristics. The matching process first identified the proximity of each NY-Sun project to a project in the supplemental data set, then checked the size and install dates of the project in close proximity, to identify any likely matches.

Of these, only three projects had a potential duplicative project within ±20% of the rated system size and a completion date within 365 days. Each of these cases was reviewed and the potential matches were found to be separate systems at a nearby address. Thus, the NY-Sun/DER impact sample and the supplemental data set are distinct, non-overlapping sets of projects.

2.1.2 NY-Sun projects in the supplemental data set

A similar matching process between the supplemental data set and the full NY-Sun tracking data set was used to indicate how many of the supplemental projects are likely to be NY-Sun projects. This review determined that the majority of the supplemental projects are likely to have received an incentive from the NY-Sun Program.

2.2 Data cleaning and screening

The NY-Sun/DER impact sample projects were reviewed as part of the prior analysis. That review included some follow-up with the data providers to resolve anomalies where possible, and the exclusion of a small number of projects with unresolved major inconsistencies, particularly recorded capacity values inconsistent with the production data. The supplemental data had a different screening process before it was provided to this study for analysis. That screening excluded projects whose summer production levels were inconsistent with the expected production level.[3]

For the present study, additional screening was conducted of both the NY-Sun/DER impact sample and the supplemental projects. The analysis for this study is based on a decay rate in irradiance-normalized monthly production, relative to a basic level that varies by calendar month due to seasonal changes in insolation. Months with 0 or partial production due to temporary disruptions would contribute to large changes down and up from one month to the next, at a level that can mask the more gradual degradation of system performance when operating. To avoid effects of such disruptions on the performance loss rate analysis, the data was screened in a few different ways, depending on the type of analysis. The analysis types and corresponding screening were as follows.

1. **Annual** analysis, used for general descriptions and displays of annual performance included the following data restrictions.
    a. **Full-year screen**: For each project, included only the production years with positive production in each of the 12 months.
    b. **Change screen**: For each project, excluded any year-to-year interval with more than a 20% change in either direction.
2. **Monthly** analysis**,** used for estimation of PLR used the following data restrictions.
    a. **Disruption exclusion.** For each project, excluded any month with zero production, as well as the preceding and following month. The exclusion of adjacent months means the months with partial production due to components or communications breaking or getting repaired mid-month do not affect the estimated performance loss rate.
    b. **Minimum series length**. For monthly analysis using separate PLR estimates for individual projects, only projects with at least 24 months of data were included. This requirement ensures that a meaningful performance loss rate can be estimated for the project, given the variation in underlying production rates due to seasonal patterns.

These screens mean that the performance loss rate analysis reflects the performance loss rate for operating, communicating projects. At the same time, the rate of lost production associated with temporary disruptions or disrepair is also of interest. This issue is addressed at a high level by reviewing the proportion of projects with disruptions identified as a function of time since installation (see Section 2.3.5, Effects of temporary disruptions). The screening is described further in Section 3.2.1, Data cleaning. Factors and expectations related to persistence over longer

[3] Specifically, the original screening prior to receipt by this study excluded all months of production for any project that had more than three months in the period April to October with actual production 1.5 times or more than the expected production for that period.

time frames than could be empirically addressed in this study are discussed in Section 2.4, Longer-term persistence.

The data used in the PLR analysis after the disruption exclusion is summarized in Table 2-1. Comparison with Table 1-3 shows that this screening excluded only a handful of the NY-Sun/DER impact sample projects, and a small percentage of the projects in the supplemental data. The earlier analyses of the NY-Sun projects included some review and corrections of anomalous cases, which may have limited the frequency of bad data in the series used. Similarly, some screening of anomalous values was applied to the supplemental data prior to it being made available to this study.

Table 2-1. Data used in the PLR analysis, after disruption exclusion

Source	Sector	Size category	Project capacity (kW)	Number of projects	Time frame of installations	Time frame of production data	Median months of production data
NY DER Portal	Non-residential	<200	55–200	30	2012–2017	2012–2024	69
NY DER Portal	Non-residential	200–750	200–743	133	2012–2018	2013–2024	83
NY DER Portal	Non-residential	≥750	783–2994	79	2013–2018	2013–2024	83
NY-Sun installation contractors	Non-residential	<200	6–189	33	2012–2016	2013–2024	59
NY-Sun installation contractors	Non-residential	200–750	200–348	6	2012–2016	2013–2024	34
NY-Sun installation contractors	Residential	<200	3–17	105	2016	2016–2022	88
DNV supplemental NYS data	Residential	<200	1–37	13,298	2012–2023	2013–2023	47

Table 2-2 summarizes the projects used for the different stages of analysis.

Table 2-2. Project counts by analysis type

	Non-residential	Residential NY-Sun	Residential supplemental	Total
Total projects	281	105	13,587	13,973
Projects in annual CF review	277	104	8,722	9,103
Projects in pooled monthly analysis	281	105	13,298	13,684
Projects with project-specific PLR (min 24 months)	277	105	8,599	8,981

There are a number of reasons for gaps in the data, as well as for discontinuation of the data for a particular project.

- A system may go out of service temporarily or permanently.
- A system may continue to operate, but communications to the system manager may be disrupted or permanently discontinued.
- A leased system may be disabled due to nonpayment.
- The service manager may go out of business. The systems may or may not be turned over to another manager.

Most of these reasons do contribute to some level of production loss, whether temporary or long-term. The production data alone cannot identify whether a disruption or early termination of the series is a lack of production or a communications problem. For this reason, the PLR analysis presented here addresses operating communicating systems only. A separate analysis of the proportions of projects with zero and near-zero production, included in Section 2.3.5, Effects of temporary disruptions, provides some insight into the potential overall loss rates including intermittent disrepair. Effects of systems going permanently offline or being abandoned cannot be assessed by the data available for this study. Section 2.3.4, Isolating effects of system characteristics on performance loss rates, includes some discussion of these effects.

2.3 Results

2.3.1 Available cases by years of observations

Table 2-3 indicates the number of projects in the analysis for each number of years since installation, by sector.

Table 2-3. Number of projects in the analysis after disruption exclusions by years of data since installation

Sector	Size (kW)	1	2	3	4	5	6	7	8	9	10	11	12
Non-residential	<200	63	63	56	49	39	19	10	6	2	2	1	0
Non-residential	200–750	139	139	135	120	99	90	66	42	24	13	3	0
Non-residential	≥750	79	79	76	74	67	60	38	17	7	2	0	0
Residential	All	11249	8976	7834	6722	5015	1208	231	88	68	52	50	47
Total	All	11530	9257	8101	6965	5220	1377	345	153	101	69	54	47

The table shows a solid set of projects with up to six years of data, and smaller numbers with higher numbers of years of data. The large number of cases with four to six years of data provide for some clear insights into how production levels change over these first several years of the projects, as described below. The drop-off in the higher years reflects both the smaller number of projects installed earlier and the challenges of continuing to receive data over many years.

2.3.2 Capacity factors over time[4]

Figure 2-1 shows the distribution of capacity factors for each number of years since system installation. The projects shown correspond to those that passed the annual screening, described in Section 2.2.

Capacity factors on the order of 13% would be expected for these projects. Multiple Year 1 impact analyses of the NY-Sun program have produced CF estimates in the range of 11% to 14.4% for that program. This is the level of the median CF values for each year, indicated by the heavy bars in Figure 2-1. Values higher than about 20% would be physically impossible given actual solar radiance patterns. Figure 2-1 indicates some projects with CF above this level, and even above 100%. A likely explanation for these unrealistically high CF values is that the system capacity on record for the project is lower than what was actually installed, so the observed production is higher than is realistic for the recorded capacity. Likewise, some of the very low CF values may reflect recorded capacity higher than what was actually installed, though these also include instances of actual low production. For the most part, the CF values at unrealistically high levels are associated with the supplemental data set; the original use of that data did not include direct screening for disparities between the recorded capacity and the observed production.[5]

---

[4] The capacity factor distributions shown in Figure 2-1 and Figure 2-2 include only project-years with a full 12 months of production data, and exclude any project-years with more than a 20% change in CF between successive years. These restrictions were applied to avoid added noise from incomplete years of data, or large swings due to temporary disruptions. As shown further in Section 2.3.5, Effects of temporary disruptions, temporary disruptions occurred very rarely in the data.

[5] The supplemental data had been screened to exclude projects with actual production deviating from expected production by more than a specified amount during the summer months. However, those data did not screen or correct for recorded capacity inconsistent with the production data.

Figure 2-1. Distribution of capacity factor by number of years elapsed since installation, screened projects

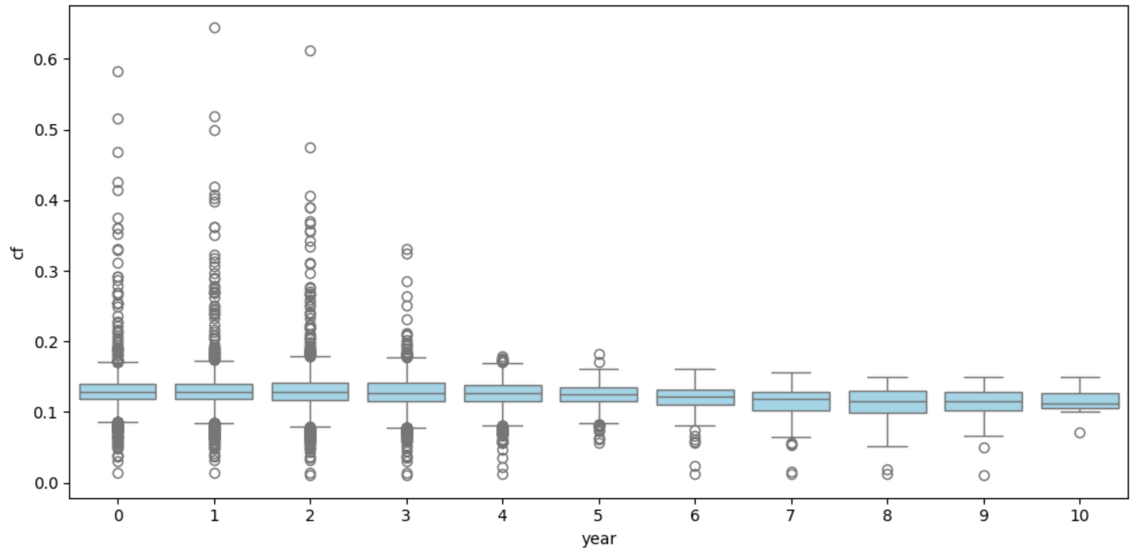
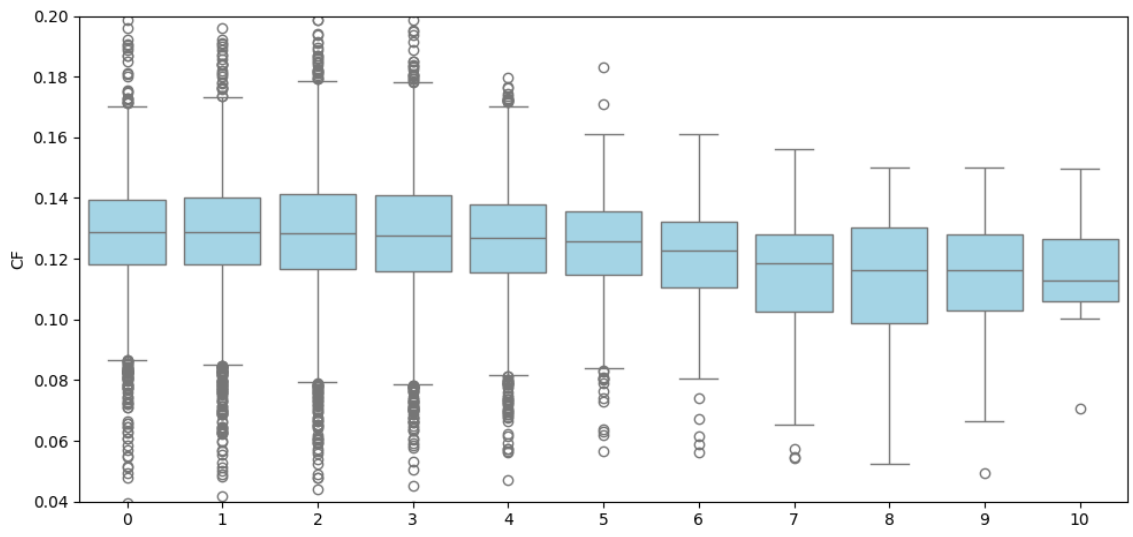


Figure 2-2 shows a close-up of the same plot, restricting to a narrower CF range that includes most of the individual values. In the figure, each box runs from the lower to the upper quartile of CF values for that year, with a dotted line at the mean and a solid line at the median. The upper and lower lines indicate bounds beyond which observations are expected to be rare. Values beyond these bounds are indicated individually.

Figure 2-2. Distribution of capacity factor by number of years elapsed since installation, screened projects (close-up)



The figure shows that the bulk of CF values across all projects and years are in the range of 11% to 15%, consistent with the prior impact evaluation findings of an overall CF around 13%.

The data screening used for the annual analysis excludes extreme values of change in year-over-year CF, and the monthly analysis excludes temporary zero months and adjacent months, but all the analyses include any projects with consistently high CF. If the application system capacity is understated, the CF would be expected to be high for all months, but the proportional loss over years or months would be the same as if the correct system capacity were used to calculate CF. Hence, even the cases of unrealistically high CF noted in Figure 2-1 are allowed in the analysis because the focus of the screening is based on variability and not on levels.

The figure also shows that the average and median CF vary only within a range of 2 percentage points over the years of observations.[6] Thus, the performance loss rate to be estimated–the year-over-year decline–is a small value, as has been found in other studies.

The experience of the evaluation contractor's solar experts who advise system developers is that the production level for a given project tends to increase between the first year and the second or third, with declines observed thereafter. The increasing production in the first couple of years is related to working out the operations and communications systems. This pattern does appear to be evident in the medians displayed in Figure 2-2. (There is a slight uptick in Year 9, but very few cases are included in this group, as indicated in Table 2-3.)

To explore the early increasing production further, the evaluation contractor team attempted to estimated PLR separately for an initial period and a later period. However, despite some indication of lower initial loss rates in the annual data, the more detailed estimation using monthly data did not identify meaningful differences in PLR for earlier and later periods.

2.3.3 Performance loss rates

2.3.3.1 How performance loss rates are calculated

As described further in Section 3.3, Performance loss rate analysis, the PLR for this study was calculated as a *proportional* loss rate. This approach assumes the same percentage reduction in

---

[6] Except for the project on year 11, where the CF is 17%.

CF across all months in the calculation span (on a seasonal-adjusted basis). This assumption is consistent with exponentially declining production.

In an annual analysis, the proportional loss rate is calculated from the average logarithm of the ratio of each year's production to the prior year's. For the monthly analysis used in this study, the analysis accounts for the fact that seasonal insolation patterns account for large, regular month-to-month changes in CF. The proportional performance loss rate can be thought of as the average proportional year-over-year decline in each month's production, relative to that for the same month in the prior year, averaged over all months and years of the study.

The PLR was calculated in three ways for this study, as described further in Section 3, Methods.

1. **Average project-level PLR, across months.** A PLR was calculated separately for each project, across the months of data available for the project. The average across projects of the Project-level PLR was then calculated by for each type within each category of interest.
2. **Pooled PLR by category, across months.** The PLR was calculated by type within a category, via regression across all projects and months in that type.
3. **Pooled PLR by category and month.** The PLR was calculated by type within a category and month of the year, via regression across all projects and years. Monthly PLRs were averaged to produce the annual PLR. Two averaging methods were used.
    a. **Simple averages of the 12 monthly PLR values**
    b. **CF-weighted average of the 12 monthly PLR values.** This averaging gives more weight to the months that contribute more to annual production.

In general, these methods gave similar results. Method 3, PLR estimates by month, are possible only with the pooled estimation method, as individual projects do not have enough observations for a single month of the year to develop meaningful PLRs by month. The project-level PLR estimates averaged by subgroup provide the most intuitive and consistent results, and are used for most of the analysis presented in this report. However, the pooled estimates by month offer additional insights and avoid some potential biases in the project-level results. The next section provides high-level results comparing the methods.

2.3.3.2 PLR overall and by region

Table 2-4 summarizes PLR overall and by region, using the average project-level (1) and pooled across-month (2) estimates. As noted, the results are similar for the two methods. Across all projects in the study pool, the PLR estimated by these methods is a little over 1% per year. The results are well determined, with standard errors one-tenth to one-quarter of the estimate.

Table 2-4. PLR estimates overall and by region (percent performance loss per year)

Category & type	All	Region – Downstate	Region - Long Island	Region - Upstate
Number of projects	8,981	1,670	4,946	2,365
Avg capacity (kW)	31	22	10	81
Median number of months	59	60	58	62
Avg. CF	12.7%	12.9%	13.2%	12.0%
1. Average project-level PLR	1.17%	0.96%	0.42%	2.89%
Std err (Average project-level PLR)	*0.10%*	*0.23%*	*0.13%*	*0.21%*
2. Pooled PLR by category across months	1.23%	0.96%	0.48%	3.00%
Std err (Pooled PLR by category across months)	*0.15%*	*0.24%*	*0.12%*	*0.12%*

In Table 2-4, the number of the projects shown in each column is the number of unique projects in the analysis for that group, each with multiple months of observations. The “median number of months” row indicates the median, across all projects in the analysis, of the number of months in the analysis. Across each of the subgroups in the table, the median is around 60 months or 5 years. The average CF for the projects in the study is 12.7%. This is the same as the first-year CF reported in the most recent impact evaluation of the NY-Sun program.

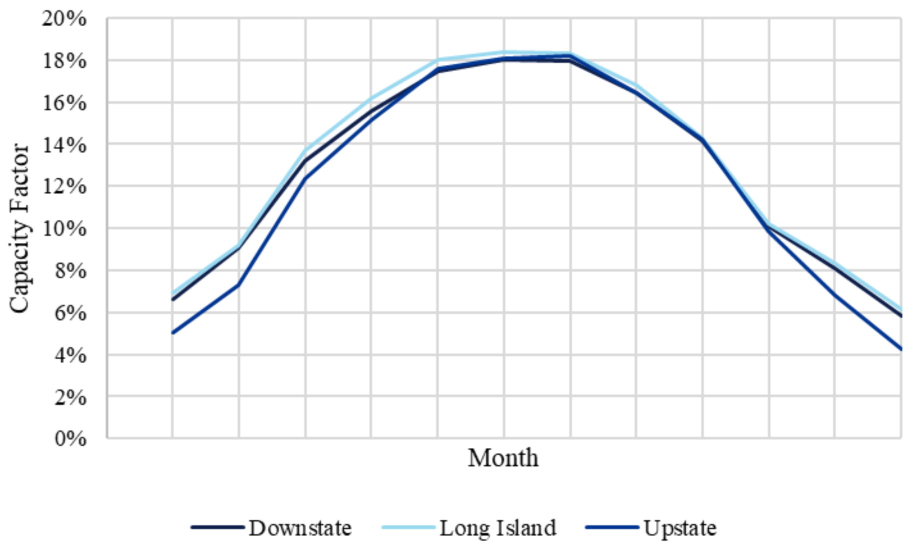
The total estimates of over 1% per year are somewhat high relative to the benchmark of 0.64% per year. The Downstate result is under 1%, and the PLR for Long Island is under 0.5% per year, on the low end of the prior studies. The PLR estimated for Upstate is substantially and significantly higher, at 2.9% per year. This high value contributes to the high overall value.

One possible reason for high Upstate PLR could be related to snowfall in the winter months. Both the average project-level PLR and pooled PLR across months in Table 2-4 estimate a single loss rate across all months of the year. To the extent PLR is systematically higher in seasons with systematically higher or lower CF, these overall calculations could over- or understate the loss rates for the year as a whole. In particular, if the PLR is very high in the winter, which contributes little to annual production, estimating a single PLR across all months could lead to an overstated annual PLR.

For this reason, the pooled PLR by month was reviewed, weighted by each month's CF (estimation method 3b). This weighting lessens the contributions of winter months, with low CF, to annual PLR.

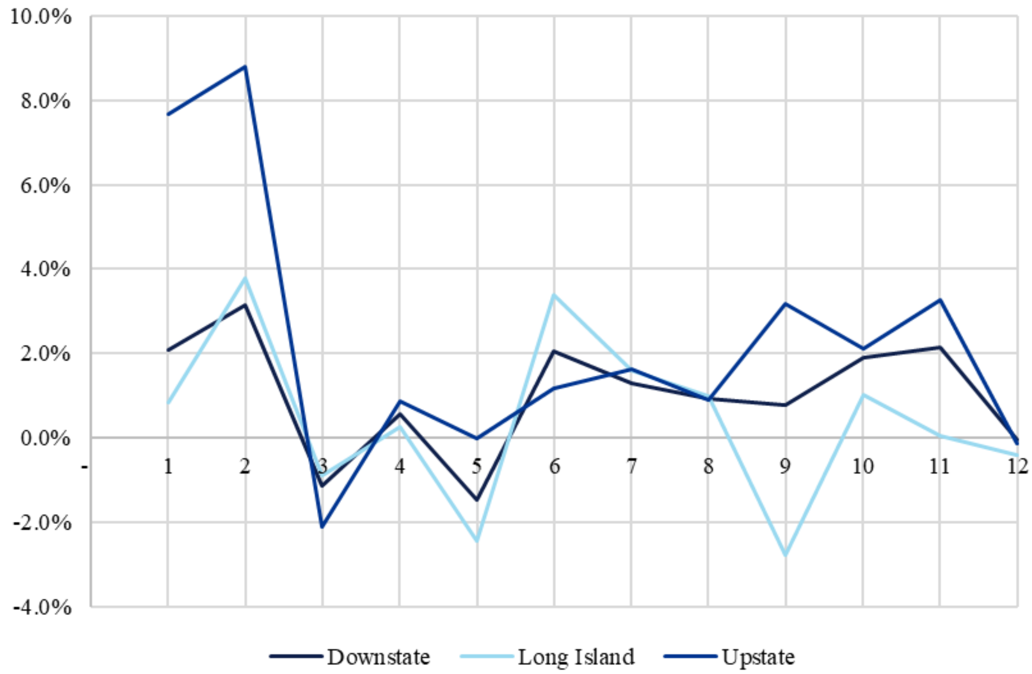
Figure 2-3 shows the average CF by month for each region. The figure shows that the average production is roughly three times as high in June and July as it is in January, for all the regions. Thus, if the PLR is relatively high in the winter but not in the summer, an estimate counting all declines equally regardless of CF level will overstate the decline in annual production.

Figure 2-3. Average capacity factor vs month of the year, by region



To investigate this possibility, the PLR was estimated separately for each month of the year, by region (the third method noted in Section 2.3.3.1, How performance loss rates are calculated). Results are displayed in Figure 2-4. The figure shows considerable bouncing around from month to month, with little systematic seasonal pattern, and averaging around the overall values indicated in Table 2-4.

Figure 2-4. PLR vs month by region (percent per year)



The figure shows an Upstate PLR particularly high in January and February, contributing to the high overall value in Table 2-4. However, when the monthly PLR is weighted by the monthly CF (which is proportional to the monthly production), the annual value is less extreme, though still higher than for the other regions. Results are shown in Table 2-5.

Table 2-5. Annual PLR (% loss per year) by region, by different methods

Region	Number of projects	Avg capa-city	Avg CF	PLR estimated across months: 1. Average project-level PLR	PLR estimate d across months: 2. Pooled PLR	3. Pooled PLR estimated by month: 3a. Simple average monthly pooled PLR	3. Pooled PLR estimated by month: 3b. CF-weighted average monthly pooled PLR	3. Pooled PLR estimated by month: 3.c SE CF-weighted average monthly pooled PLR
Downstate	1,670	22	12.9%	0.96%	0.96%	1.02%	0.87%	0.19%
Long Island	4,946	10	13.2%	0.42%	0.48%	0.45%	0.41%	0.13%
Upstate	2,365	81	12.0%	2.89%	3.00%	2.27%	1.67%	0.15%
All	8,981	31	12.7%	1.16%	1.23%	1.04%	0.83%	0.09%

The first two columns of this table are the same as the values shown in Table 2-4. The last two are based on the pooled PLR estimates by month.

For Downstate and Long Island, the results are fairly similar across all four estimation methods. For Upstate, however, the CF-weighted average of the monthly PLRs is much lower. While this performance loss rate of 1.7% per year is still high compared to the other regions and to prior studies, it is less extreme than the PLR estimated across all months jointly (methods 1 and 2), or the unweighted average of the monthly PLRs (method 3a).

A possible reason for high PLR in the Upstate region is that heavy snowfall contributes to system damage at higher rates. A possible additional contributor could be heavier snowfall in later years than in earlier years for the particular arc covered by most of the data for the Upstate projects. The irradiance normalization used in this study adjusts all the CF observations to long-term average cloud cover conditions but does not adjust for snowfall. Snowfall adjustments are more complex.

One study reviewed found higher loss rates associated with warmer climates, in contradiction to these regional results.[7] However, that study was comparing across different states in more southern latitudes, and did not involve regions with higher and lower snow cover, as in New York State.

The team regards the CF-weighted averages of the monthly pooled PLR in the final column of the table as the best PLR estimates based on this analysis for the results at the regional and overall levels. The results in the final row of Table 2-5 are the averages of the regional results, weighted by the number of projects in each region.

If the estimated 0.83% annual performance loss rate were to persist through the system lifetimes, New York could still achieve approximately 80% of the expected annual energy production by 2050 from the 10 GW of solar capacity planned for installation by 2030 under the Climate Act Scoping Plan. The production loss by 2050 would be on the order of 2,200 GWh per year, out of over 11,000 GWh of first-year production, or 20%.

The pooled PLR model estimated across months includes:

- A term for each calendar month, accounting for the seasonal variation in production due to solar radiance

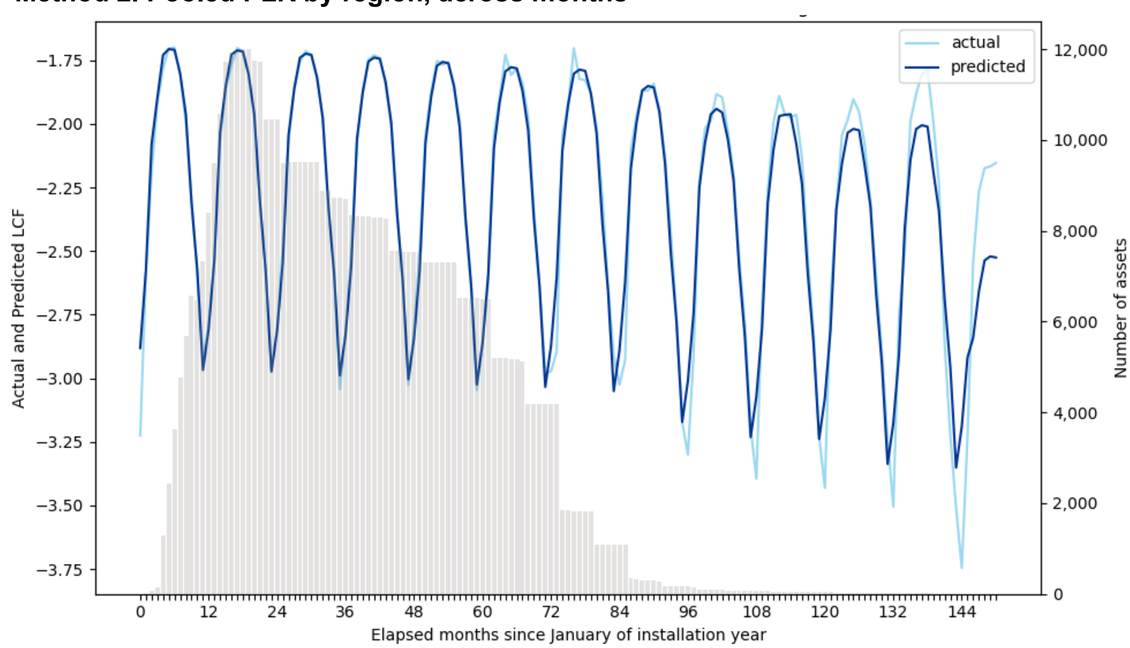
[7] Source: https://www.nrel.gov/pv/fleet-performance-data-initiative.html

- A term for each project, accounting for variation across projects in factors that affect the project's overall production level across months of the year.
- A trend term for each region, accounting for declining production over time, controlling for the calendar month pattern and the project-specific overall production levels.

The model is estimated on a log-linear basis. That is, the log of CF is modeled as the sum of the calendar month term, the project-specific level term, and product of the trend term and the number of months since installation. The trend term is the PLR.

The pooled PLR model fit for all regions is illustrated in Figure 2-5. The heavy blue line is the average model fit for each month, given by the month effect, the region-specific trend, and the average of the project effects across all projects with data in that month. The light blue line is the average of the actual observations across the same projects. Also indicated are the number of projects contributing to the regression in each month.

Figure 2-5. Actual vs predicted log of CF and number of projects – all regions
Method 2. Pooled PLR by region, across months



Across the first five to six years, the model fit tracks actual CF very closely. After that point, the fit is less tight. There also are very few projects with data this far out. Through the more stable first five to six years, the production cycle each year is clearly visible in both the actual and fitted data, along with a general declining trend. The more erratic ups and downs in the later years reflect the effects of different projects being included; because of the project-specific term in the model, the average fit continues to track the average actual usage even in these years when the group of projects being averaged is different month to month. Because there are very few observations in the later period, these points have little effect on the overall fit and PLR estimate.

Figure 2-6, Figure 2-7, and Figure 2-8 show the same model fits and actual values as in Figure 2-5, broken out into each region's projects. The same monthly patterns are visible in all three, with somewhat different levels based on the projects for each region, and the region-specific trends. The downward trend is most pronounced for the Upstate fit, and quite mild for the other two regions, consistent with the PLR values shown in Table 2-5. The average actuals (light blue) in the Upstate fit also shows more extreme low CF in the winter and high CF in the summer relative to the modeled overall monthly factors (in dark blue) that are dominated by the other two regions in the overall fit shown in Figure 2-5.

Figure 2-6. Actual vs predicted log of CF and number of projects – Downstate Method 2. Pooled PLR by region, across months

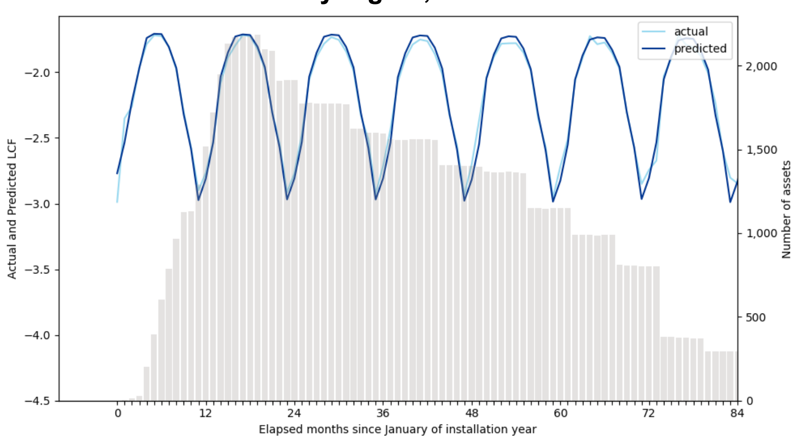


Figure 2-7. Actual vs predicted log of CF and number of projects – Long Island Method 2. Pooled PLR by region, across months

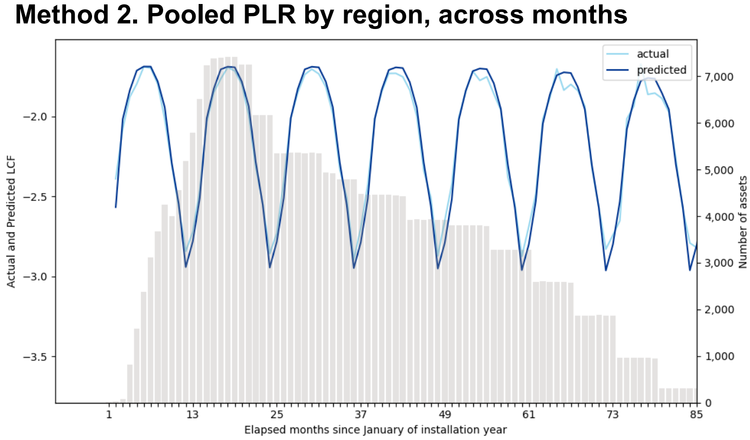
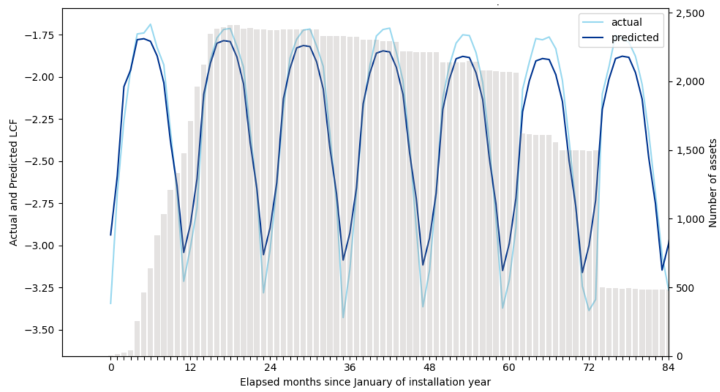


Figure 2-8. Actual vs predicted log of CF and number of projects – Upstate
Method 2. Pooled PLR by category (region), across months



To examine PLR by other subgroups, the evaluation contractor team used the PLR calculated separately for each project, averaged by subgroup (Method 1 from Section 2.3.3.1, How performance loss rates are calculated). This is the most direct way to examine PLR by subgroup. However, this approach means that differences across types within a particular category may be due to different mixes of other factors. In particular, given the Upstate results discussed here, differences in the mix across regions could lead to differences by subgroup that are not directly related to the subgroup type. Therefore, the project-specific average PLRs by subgroups are for informational purposes while the overall pooled PLR of 0.83% is representative of effective PLR. The cross-factor analysis allows for further understanding of the trends within a subgroup.

2.3.3.3 PLR by sector

Table 2-6 indicates the study findings across all projects with sufficient data for project-level PLR calculation by sector. All the non-residential projects are from the NY-Sun/DER sample. Residential projects are from the NY-Sun/DER sample and the supplemental NYS data set. First-year savings for the NY-Sun/DER sample were estimated in prior impact studies of the program.[8]

---

[8] Solar Photovoltaic Program Impact Evaluation for 2008 and 2011-2016 and NY-Sun Solar Photovoltaic Program Impact Evaluation for May 1, 2016, through March 31, 2018, NYSERDA prepared by DNV GL.

Table 2-6. PLR by sector (percent production loss per year), Method 1. Average project-level PLR, across months

Category	Number of projects	Avg. capacity (kW)	Median number of months	Avg. CF	Average project-level PLR	SE
All	8,981	31	59	13%	1.17%	0.10%
Residential	8,704	7	59	13%	1.17%	0.10%
Non-residential	277	792	72	12%	1.02%	0.41%
Non-residential, <200 kW	62	103	57	12%	0.45%	0.96%
Non-residential, 200 – 750 kW	136	423	77	12%	1.30%	0.65%
Non-residential, ≥750	79	1970	79	12%	0.98%	0.46%

The first line shows a PLR of 1.2% per year for all projects in the study data set, consistent with Table 2-4. The Residential sector, with the large majority of projects, has similar results to the overall. The small non-residential projects show a lower PLR, but with a high standard error, so that it is not statistically significantly different from the other non-residential size groups. The lower PLR for the smaller projects in part may be related to a much lower proportion of Upstate projects and higher proportion in Long Island compared to the other non-residential systems. The medium and large non-residential projects have higher PLR, more similar to that for the residential sector overall. The residential and non-residential sector-level PLRs are not statistically significantly different.

The differences among the non-residential size groups are not statistically significant due to the relatively small number of non-residential projects in total. The standard error that determines statistical significance is a function of the number of projects and the variability in year-over-year changes. The non-residential results are based on many fewer projects than the residential results, hence, they tend to have higher standard errors.

2.3.3.4 PLR by system characteristics

Table 2-8 presents the average project-level PLR by additional subgroups, similar to Table 2-4 and Table 2-6. The regional results were discussed in Section 2.3.3.2, PLR overall and by region, and results by data source, sector, and size were discussed in Section 2.3.3.3, PLR by data source and sector. Results by other categories of interest are discussed below.

Table 2-7. PLR by subgroup (percent production loss per year), Method 1. Average project-level PLR, across months

Category	Type	Number of projects	Avg capacity (kW)	Median number of months	Avg. CF	Mean PLR	Std err
All	All	8981	31	59	12.7%	1.17%	0.10%
Region	Downstate	1670	22	60	12.9%	0.96%	0.23%
Region	Long Island	4946	10	58	13.2%	0.42%	0.13%
Region	Upstate	2365	81	62	12.0%	2.89%	0.21%
Sector	Non-residential	277	792	72	11.8%	1.02%	0.41%
Sector	Residential	8704	7	59	13.0%	1.18%	0.10%
Size	< 200 kW	8766	7	59	13.0%	1.17%	0.10%
Size	200 - 750 kW	136	423	76.5	11.7%	1.30%	0.65%
Size	> 750 kW	79	1970	79	11.9%	0.98%	0.46%
DAC	NonDAC	6040	32	60	12.8%	1.25%	0.12%
DAC	DAC	2941	29	58	13.1%	1.01%	0.18%
Purchase Type	Lease	7337	7	60	12.8%	0.91%	0.10%
Purchase Type	PPA	1380	82	49	13.2%	2.63%	0.35%
Purchase Type	Purchase	88	380	57.5	12.0%	2.69%	0.74%
Purchase Type	Unknown	176	461	59	11.9%	-0.19%	0.71%
Module Family	Monocrystalline	3031	7	42	13.2%	-0.12%	0.19%
Module Family	Polycrystalline	2131	8	61	13.1%	1.95%	0.19%
Module Family	Unknown	3819	63	60	12.4%	1.76%	0.15%
Microinverter	Microinverter	6750	7	60	13.0%	1.46%	0.11%
Microinverter	Optimizer	1675	7	51	12.7%	0.17%	0.30%
Microinverter	String	272	182	73	12.1%	0.21%	0.50%
Microinverter	Unknown	284	606	67	11.9%	1.00%	0.38%
Inverter Manufacturer	G	6818	7	60	13.0%	1.46%	0.11%
Inverter Manufacturer	H - Other	185	269	68	12.2%	0.03%	0.78%
Inverter Manufacturer	I	1676	8	51	12.7%	0.18%	0.30%
Inverter Manufacturer	Unknown	302	566	78	11.9%	0.90%	0.27%
Panel Manufacturer	A	1637	14	62	12.8%	1.64%	0.11%
Panel Manufacturer	B	1504	10	44.5	13.3%	-0.05%	0.36%
Panel Manufacturer	C	374	7	57	13.3%	2.04%	0.52%
Panel Manufacturer	D	2349	7	55	12.8%	0.77%	0.17%
Panel Manufacturer	E - Other	1609	80	66	12.3%	1.61%	0.25%

Category	Type	Number of projects	Avg capacity (kW)	Median number of months	Avg. CF	Mean PLR	Std err
Panel Manufacturer	F	1427	10	61	12.9%	1.89%	0.26%
Panel Manufacturer	Unknown	81	981	65	11.9%	0.37%	0.57

2.3.3.5 PLR by DAC

Table 2-8 indicates a significantly lower PLR for DAC than non-DAC geographies. This is likely related to their regions. In the study pool, 82% of the DAC projects are in Long Island, where the PLR is low, and only 7% are Upstate, where the PLR is high. Non-DAC projects by contrast are 42% in Long Island and 36% Upstate.

2.3.3.6 PLR by purchase type

The large majority of projects in the study data were leased. In Table 2-8, these projects show PLRs close to those seen for the overall, as expected given that they are the majority of this total. PPA systems exhibit much higher PLR than leased. Purchased systems are relatively few in the study data and have much higher PLR than leased. This observation aligns with anecdotal evidence from system owners. Leased systems typically come with performance guarantees, allowing owners to address underperformance issues promptly. In contrast, Power Purchase Agreement (PPA) systems often lack such guarantees, leading to longer periods of underperformance before issues are resolved.

Purchased systems in the residential sector are likely to have less consistent maintenance, since owners may not realize the systems are not performing well and may not call for repairs immediately. Thus, the higher PLR for these systems makes sense.

While the study cases are heavily weighted towards leased projects, the NY-Sun Program has had about 27% leased, 10% PPA, and 63% purchased systems since 2012.

2.3.3.7 PLR by module family

Monocrystalline systems exhibit very low PLR, not statistically different from 0. By contrast, polycrystalline systems exhibit a PLR somewhat higher than the overall average. The polycrystalline systems are generally older, and there may be reasons other than the module family that they have higher performance loss rates.

Over one-third of the systems in the study data have an unknown module type. The majority of these are older systems and likely have a mix of module types. The projects with unknown module types have a PLR similar to the polycrystalline systems.

2.3.3.8 PLR by inverter type

Microinverters have higher PLR than optimizer or string inverters. The lower PLR for string inverters in particular is somewhat counterintuitive. String inverters tend to be used for larger loads, as indicated by the much higher average capacity shown in Table 2-8. The heat from high loads can lead to component stress and failure. On the other hand, the larger systems associated with string inverters may have better maintenance.

Perhaps more importantly, microinverters can fail one by one within a system, and the owner, especially a residential owner, may not notice production tailing off. By contrast, if a string inverter has failed, the production will go to zero. Leased systems and systems owned by homeowners are both at risk of having one or two inverters fail and not having them replaced in early years. Homeowners may not notice, and third-party operators may not care if the system is not going to be violating its performance guarantee. Section 2.3.4, Isolating effects of system characteristics on performance loss rates, shows that even when controlling for other project characteristics, the microinverter performance loss rate is higher than that for string or optimizers.

A question of interest is whether the higher PLR for microinverter systems is associated more with older systems, and whether it applies across all purchase types. However, the available data doesn’t support comparing PLR at that level of granularity.

2.3.3.9 PLR by component manufacturer

The PLR results in Table 2-8 vary across panel manufacturers. Manufacturers A, C, F, and Other all have PLR over 1.5% per year. Manufacturer D has PLR more in line with the overall average, at 0.8% per year. Manufacturer B is distinguished by a PLR not statistically significantly from zero. Notably, the median span of data for this manufacturer is less than four years. Thus, these lower loss rates could be related to improved performance in early months, with insufficient time to see that offset, or could be related to having more recent technologies installed.

Table 2-8 also shows results by inverter manufacturer. The average capacity indicates that Manufacturers G and I have mostly residential and small commercial projects. Results for Manufacturer G are consistent with overall residential results shown earlier. Manufacturer I shows a much lower loss rate, under 0.2% per year. This manufacturer also has few median months of data, indicating more recent equipment.

The Other group, with a small number of cases, has more large non-residential projects. Unlike the general non-residential results, Manufacturer H has a very low PLR, but it is not statistically significantly different from the overall non-residential result.

Given the number of factors that could be associated with manufacturers other than the equipment itself, it is difficult to draw conclusions from these results alone. The regression results presented in Section 2.3.4, Isolating effects of system characteristics on performance loss rates, attempt to isolate the effects of the various factors.

2.3.4 Isolating effects of system characteristics on performance loss rates

2.3.4.1 Overview and interpretation of the cross-factor PLR analysis

The comparisons of PLR by individual characteristics in the tables above are informative, and in many cases point to differences that make sense. At the same time, some of the results show patterns that are at odds with what might be expected, or they have no clear pattern.

The PLR differences by subgroup reflect not just the effect of the main characteristics displayed in each table, but also the effects of other characteristics that may be different between the groups being compared. Thus, assessing the PLR differences associated with a given characteristic requires analysis that isolates the effects associated with each characteristic.

To attempt to address the potential for confounding of factors affecting PLR, the next step in the analysis was a regression analysis that included multiple project characteristics at once. This analysis estimated how each characteristic is related to higher or lower PLR *while holding the other characteristics fixed*.

The cross-factor regression analysis was conducted for all the projects for which a project-specific PLR was estimated. These were projects with at least 24 months of production data.

As described further below, the regression results confirm most of the directional relationships indicated by the single-characteristic comparisons presented in Section 2.3.3, Performance loss rates. They also support the importance of controlling for regional differences when making comparisons while also providing indications that many of these comparisons are driven by other associated characteristics.

The results of the cross-factor analysis are not intended as the basis for estimating the effects of particular features. Rather, these results confirm, and in some cases refine, qualitative relationships identified from the comparisons of subgroup averages.

2.3.4.2 Cross-factor analysis results

The cross-factor analysis determines the incremental addition or subtraction to the PLR associated with each characteristic. Along with the estimated increments, the analysis provides a standard error of that estimate. In technical terms, the analysis is a multiple variable regression, the increments are the coefficients for each factor in the regression, and the base quantity is the regression intercept.

Results are shown in Table 2-9. An initial version of this analysis included all the characteristics displayed in Table 2-8. For several of these variables, the estimated incremental effects were not statistically significant. The analysis presented in Table 2-9 includes only characteristics with statistically significant incremental effects.

The results in Table 2-9 include the base factor and the incremental additions to the PLR associated with each characteristic. For each characteristic category, the analysis sets the first type in the category as the neutral level, and identifies the incremental effect of each other type relative to that neutral. That is, the non-zero increment for a given category type represents the difference in PLR between that type and the neutral category type, all other characteristics held constant. The neutral characteristics are indicated in the table by light gray highlight.

For any one project in the analysis, the estimate of its PLR based on the analysis is the sum of the increments for that project's characteristics, including a base quantity that is included for all projects. Thus, for a project with all the "neutral" characteristics, whose incremental contributions are all 0, the analysis estimates the PLR as just the base value 0.07% per year. For a project with all the most common characteristics (by project count)—Long Island, non-DAC, Leased, Unknown Module family, and Microinverter inverter type—the estimated PLR from this analysis would be the sum of the increments for each of these characteristics, plus the base. Thus, the estimated average PLR for projects of those characteristics would be 0.67% per year. The average estimate from the regression across all the projects used to fit it is 1.2%, the same as the average PLR of these projects. (This match of the average predicted value to the average observed value is a structural feature of the analysis method.)

These examples are to illustrate the meaning of the increments listed. However, this analysis is not intended to generate predictive estimates for particular combinations of characteristics. Rather, the purpose of the analysis is to identify which factors appear to be associated with higher or lower PLR, holding other factors constant.

Table 2-8. PLR increment associated with each project characteristic holding other characteristics constant

						PLR increment (% per year)	
Characteristic	**Type**	**Number of projects**	**Avg. capacity (kW)**	**Median number of months**	**Avg. CF**	**Coeff**	**p-value**
Intercept	All	8,981	31	59	12.68%	0.07%	0.798
Region	Downstate	1,670	22	60	12.86%	0.00%	N/A
Region	Long Island	4,946	10	58	13.21%	-0.84%	0.009
Region	Upstate	2,365	81	62	12.00%	1.80%	0
DAC/non-Dac	non-DAC	6,040	32	60	12.83%	0.00%	N/A
DAC/non-Dac	DAC	2,941	29	58	13.13%	0.48%	0.064
Purchase type	Lease	7,337	7	60	12.84%	0.00%	N/A
Purchase type	PPA	1,380	82	49	13.17%	2.28%	0
Purchase type	Purchase	88	380	58	11.99%	3.36%	0
Purchase type	Unknown	176	461	59	11.94%	0.36%	0.703
Module family	Monocrystalline	3,031	7	42	13.15%	0.00%	N/A
Module family	Polycrystalline	2,131	8	61	13.10%	1.20%	0
Module family	Unknown	3,819	63	60	12.39%	1.44%	0
Inverter Type	Microinverter	6,750	7	60	13.01%	0.00%	N/A
Inverter Type	Optimizer	1,675	7	51	12.74%	-0.96%	0.002
Inverter Type	String	272	182	73	12.09%	-2.52%	0
Inverter Type	Unknown	284	566	78	11.87%	-3.12%	0

Notes:

The results are for all projects with a project-level PLR estimate.

Light grey rows are the default characteristic; coefficients are the increment to PLR associated with each characteristic relative to the default.

Dark gray cells indicate rows with coefficients not statistically significantly different from 0 at the 10 percent significance level.

In the table, the p-value indicates the smallest significance level at which the increment is statistically significantly different from 0. A p-value smaller than 0.1 indicates a PLR increment statistically significantly different from the neutral type at the 10% significance level, other characteristics held constant. Increments with p-value smaller than 0.1 are considered to be statistically significant. Increments with higher p-values are not statistically significant and are marked out by dark gray highlight in the table.

By region, Downstate is set as the neutral. Consistent with the exploration of region effects in Section 2.3.3.2, PLR overall and by region, the analysis indicates that being in Long Island is associated with much lower PLR and being in Upstate is associated with much higher PLR compared to Downstate. The differences are generally in line with those seen in the simple comparison of mean PLR in Table 2-5 or Table 2-8. Thus, the multiple factor analysis indicates that little of the direct comparisons of PLR shown in those tables are linked to system characteristics other than location.

By DAC status, Table 2-8 indicated lower PLR in DAC zones compared to non-DAC. The discussion in 2.3.3.5, PLR by DAC, suggested that the difference and direction were related to much higher proportion of DAC projects being in Long Island. Table 2-9 shows that indeed, once region is accounted for, the incremental effect of being in a DAC zone is an increase in PLR on the order of an additional 0.5% per year.

By purchase type, both PPA and purchased systems are associated with a much higher PLR than Leased. These findings are consistent with the results in Table 2-8.

By module family, polycrystalline modules are associated with a PLR substantially higher than monocrystalline, consistent with the results in Table 2-8.

By inverter type, the regression indicates that string types are associated with 2.5% per year lower PLR compared to microinverters, and optimizers with a 1% per year lower PLR. These results are directionally consistent with and somewhat stronger than that from Table 2-8. As noted in the discussion of those results, the higher PLR for microinverters may reflect degradation due to unnoticed and unaddressed failures of individual microinverters within a system.

By panel manufacturer, no statistically significant differences were identified in the initial regression, and this characteristic is not included in the analysis shown in Table 2-9.

By inverter manufacturer, a version of the analysis including this variable produced anomalous values, both for one large manufacturer and for another factor in the analysis. The evaluation contractor team determined that the manufacturer effects were likely confounded with other system effects, and excluded this factor from the analysis presented in Table 2-9.

Overall interpretation of the results. As noted, the cross-factor analysis results are not recommended as predictions of the effects of particular combinations of features. The results reinforce the importance of considering what other differences may be affecting any comparison across a particular category of characteristics and provide some indications of how those factors may be confounded.

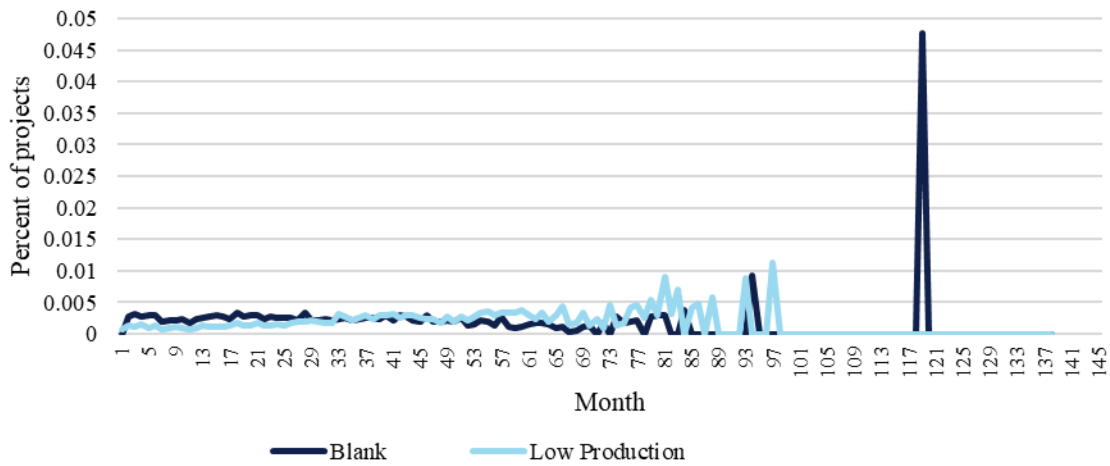
In general, the cross-factor analysis confirms the indications from Section 2.3.3.4, PLR by system characteristics. These indications include that the higher mean PLR in non-DAC zones compared to DAC is related to the regional mix in these two groups. With region controlled for, being in a DAC is associated with higher PLR, not lower. For other characteristics, the cross-factor analysis generally supports the directional patterns indicated by the simple comparison of means.

2.3.5 Effects of temporary disruptions

As noted, for each project the analysis excludes months with 0 production, as well as the adjacent months. This exclusion avoids most temporary disruptions lasting more than a month. Disruptions might be of communications or of system operations. The exclusion was applied to reduce the potential for large swings in CF that would add a lot of variability to the PLR estimates.

To assess whether there might be an additional source of increased production loss related to increasing frequency of outages, the evaluation contractor team reviewed the proportion of projects that had 0 or low production, as a function of system age in months. Low production for a month is defined as less than 5 kWh and greater than zero. For each month, a project is included in the base of the percentage only if it has positive production data for later months. The number of projects included drops over time. Results are shown in Figure 2-9.

Figure 2-9. Percent of reporting projects with 0 or low production versus months since installation



Note: Low production for a month is less than 5kWh and greater than 0.

The figure shows that for any age, around 0.5% of all projects have zero production (blank) and a similar fraction have low production, for a given month. The percentages become more erratic in later months, as there are fewer projects of that age, but there does not appear to be any overall increase in the rate of disruptions over time. The spike month 119 is due to a single project with 0 production in that month, with only 20 projects having data for that many months out. Thus, effects of temporary disruptions are not a major omission in the PLR analysis. What remains unaddressed by the analysis is the rate of full system shut-downs or abandonment over time. This is likely to be more of an issue in later years than over the relatively early years examined in this study.

2.4 Longer-term persistence

The analysis presented above is based primarily on projects with five to seven years of production data. These empirical results give a useful look at how well solar system performance holds up over time, but provide little information on what happens later in the life of a system. In this section, the evaluation team offers perspectives on what can be expected over solar system life, based on the evaluation contractor’s experience as a consultant to system developers.

Overall system life. The evaluation contractor team expects residential PV systems to be able to last for 30 years with proper maintenance. While residential PV systems don't have moving parts, they are not immune to component failure. Modules, string inverters, microinverters, optimizers, and communications hardware can all fail with differing levels of impact to system production. The factors described below all suggest reasons that production loss rates could be expected to increase at system ages beyond those captured in the analysis in this study.

String inverter failures. The evaluation contractor team expects most string inverters to have a useful life of 10 to 15 years. Once a single inverter in a string system fails, the system (if it has only one string inverter) or that portion of the system (if it has multiple string inverters) is unable to convert DC electricity to AC electricity for delivery to loads and the grid, and is out of commission.

In systems that are owned by third parties (third-party owned, or TPO), the owner has a financial incentive to replace failed string inverters so the owner can continue collecting PPA or lease payments from the homeowner. The evaluation contractor team would expect that most TPO systems with failed string inverters will have inverters replaced fairly quickly in order to resume operation. For systems that are owned fully by the homeowner, it is possible that if a string inverter fails and is outside of the warranty period, which is typically about 10 to 12 years for string inverters, the homeowner may not opt to purchase a new string inverter at cost to them. The evaluation contractor team does not have data to inform a reasonable rate of inverter replacement for cash/loan systems but expects that a portion of homeowners will not repurchase a replacement inverter to replace an inverter that has failed outside of its warranty.

Thus, some portion of purchased systems with string inverters are likely to become inoperable after roughly 10 to 15 years. For the most part, this effect is not reflected in the data used in this study, and permanently abandoned systems do not contribute to the estimated performance loss rates.

Microinverter and optimizer failures. Failures of microinverters and optimizers have a lesser impact on system performance since they are installed at the module level, and module-level power electronics (MLPE) systems are designed to bypass failures and maintain the optimal string performance. Enphase and SolarEdge are the leading providers of microinverters and optimizers in the US market respectively. Both have 25-year warranties on microinverters and optimizers. Failure of MLPEs will reduce the output of a PV system proportionally. For example, if a system has 10 modules, and one microinverter or optimizer fails, the system's output will be reduced by approximately 10%. In TPO systems, owners may or may not decide to replace a single failed MLPE component. The decision is likely dependent on the impact on revenue in a PPA, the ability of a technician to get to the site, and the potential impact of the reduced performance on a system's performance guarantee level.

In typical residential PPAs, the revenue an owner earns is directly related to the output of the system. Replacing a failed MLPE component will typically come at a cost to the system owner even if the component is under warranty, so owners may balance the cost with potential recovered production.

In leased systems, revenue is not directly correlated to energy production. Lease contracts typically have a performance guarantee set between 85% and 95% of a system's forecast performance. On larger systems, if a single MLPE component fails, the performance of the system may only reduce by 5%. If performance guarantees are set conservatively, say at 85% of forecast production, the system owner may not elect to replace the failed MLPE component since they are still able to collect 100% of lease payments until the system performs below the performance guarantee threshold for a true-up period, which can be one to three years. In this case, the owner may wait for additional MLPE components or modules to fail before rolling a truck to replace.

Communication failures. Separately, lease contracts often allow for "billing on estimates" in cases where actual system production is not being communicated to the inverter provider's environment. It is best practice to replace failed communications equipment; however, this is not always done. When a system is not communicating, it is impossible to know whether the issue is specific to communications, in which case a system may be producing at normal levels, or if the system has failed in some capacity and is not producing. Given the structure of lease contracts, it is possible that failed systems can be billed on estimates and there may be a lag between when a system fails and when the system is repaired.

Roof repairs. It is best practice for residential PV installers to evaluate a home's roof prior to installing PV. The evaluation contractor team would expect installers to install PV systems on roofs that are no older than 10 years old. Otherwise, a reroof during the useful life of the PV system is nearly guaranteed. This requires the modules, racking, MLPEs and any conduit and wiring to be fully removed so the roof material can be replaced. Modules, racking, MLPEs, and conduit would then need to be reinstalled. Removal and reinstallation are typically at cost to the homeowner even in TPO systems. If a homeowner elected not to reinstall a TPO system, they would typically be required to buy the system from the owner at the fair market value of the system. Depending on the age of the system, this may be more or less than the cost to reinstall the system on a new roof. The likelihood that a homeowner would elect not to reinstall due to cost is unlikely in younger systems, but possible in older systems. Homeowners that own their systems may also elect to upgrade technology if newer technology exists at the time.

Shading. When residential PV systems are initially modeled, the site shade scene is captured either with onsite technology like a SunEye or by using 3D imagery and satellite data. Energy forecasts assume the shade scene will not change. In reality, trees, which are major factors in residential PV shade scenes, will grow. In TPO contracts, it is the homeowner's responsibility to maintain the original shade scene. In practice, there is likely foliage growth that is not managed by the homeowner. Increased shading on the PV array will decrease production. Systems with MLPE are better equipped to handle partial shading, whereas PV systems with string inverters are highly sensitive to shade: a small amount of shading can lead to a material reduction in output. It is likely that over the 20- to 30-year life of a residential PV system, the shade scene will change slightly which can impact production, usually negatively.

Effects of changes in shading are reflected in the performance loss rates estimated in this study. It is possible that shading could become worse at greater ages, since changes in shading farther out are harder to project at the time of the site assessment.

2.5 Findings and recommendations

2.5.1 Considerations for using findings from this study

This study has explored performance loss rates for existing distributed solar projects in New York State from multiple perspectives, using data from thousands of projects. While many factors can affect performance loss rates, this study provides valuable insights into many, though not all, of these factors. The following are considerations when interpreting the results:

1. **Project age:** The majority of projects in the study had four to seven years of production data. Performance loss rates may be different as projects age beyond this range.

2. **Technology evolution:** The results are necessarily based on older systems. Newer systems generally incorporate more advanced technologies, which could lead to different performance patterns.
3. **Study data characteristics:** The overall results reflect the mix of projects and characteristics in the study data. These projects were not randomly sampled from all New York State solar projects but were selected from available existing datasets. The mix of locations and technical characteristics of the studied projects differs from those incentivized by NY-Sun. Additionally, maintenance practices for projects outside the study group might differ from those within the study.
4. **Cross-factor analysis:** The cross-factor analysis estimates the effect of each factor while controlling for others. This analysis assumes simple incremental effects of individual factors, though more complex interactive effects may exist. Furthermore, this type of analysis cannot definitively identify driving factors when multiple factors tend to occur together. For example, it is challenging to separate the effects of equipment vintage from the effects of technology types associated with particular vintages.

Given these considerations, the findings from this study should be viewed as informative but not necessarily quantitatively definitive for NY-Sun projects. The findings below emphasize directional trends over precise quantitative differences.

2.5.2 Finding 1

Performance loss rates estimated in this study represent the performance loss of operating and communicating systems, over roughly the first four to seven years of life. Temporary outages of systems or communications in principle could contribute to additional production loss across the NYS solar fleet. The frequency of such disruptions across the data and time span studied is quite low, and does not suggest the need to consider temporary disruptions as an additional loss factor over this time horizon.

Recommendation 1:

Consider a study to investigate the proportions of incentivized systems that are still in operation at different ages while considering technical features (e.g., whether they are monocrystalline or polycrystalline or tracked vs. fixed systems). Such a study could include working with the system managers to identify rates of data loss and reasons.

NYSERDA Response to Recommendation: Pending. NYSERDA will consider this study as part of future performance analyses.

2.5.3 Finding 2

Overall performance loss rates are estimated at 0.83% per year across projects in New York State. This rate is statistically significantly higher than the benchmark of 0.64% generally used by the contractor for project pool assessments on behalf of developers. If the current loss rate were to persist through the system lifetimes, New York could still achieve approximately 80% of the expected annual energy production by 2050 from the 10 GW of solar capacity planned for installation by 2030 under the Climate Act Scoping Plan.

The estimated loss rate varies substantially across regions, with lower loss rates in Long Island and higher rates in Upstate New York. These regional differences may be related to differences in weather conditions that may cause damage to equipment or differences in unaccounted snow accumulations.

Recommendation 2:

Take estimated performance loss rates by region into account when projecting future production of NYS solar systems as a whole.

NYSERDA Response to Recommendation: Pending. NYSERDA will consider the estimated performance loss rates when projecting future production of NYS solar systems as a whole

2.5.4 Finding 3

Both purchased systems and to a lesser extent PPA systems have higher PLR than leased systems. This result is expected for purchased systems, where maintenance may be worse due to less active monitoring. PPA and leased systems are typically maintained by third parties who are responsible for the performance of the systems.

Recommendation 3:

Consider offerings that could support improved maintenance for purchased systems.

NYSERDA Response to Recommendation: Pending based on continued NYSERDA program funding

2.5.5 Finding 4

Monocrystalline modules are associated with lower performance loss rates than polycrystalline. These empirical findings are consistent with expectations for these technologies. Newer systems are more likely to be monocrystalline which have a higher persistence in production over the life of the system.

Recommendation:

No recommendation, given that most new installations will be monocrystalline.

NYSERDA Response to Recommendation: N/A

2.5.6 Finding 5

Microinverter systems have higher performance loss rates than those with optimizers or string systems, even controlling for other characteristics. A possible explanation is that failures of individual microinverters within a system go undetected and unaddressed resulting in higher performance loss rates. In contrast, a string system failure would result in zero production and could be more likely to be recognized and repaired.

Recommendation 5:

Consider offerings that would support system owners identifying and addressing reduced performance that may be related to microinverter failures. One option could be to encourage continued sharing of production data with NYSERDA, in exchange for NYSERDA sending alerts for indications of maintenance needs. Such an approach could provide NYSERDA with an expanded data series for future persistence studies.

NYSERDA Response to Recommendation: Rejected. NYSERDA does not have the ability to implement this recommendation.

2.5.7 Finding 6

Controlling for other characteristics, systems in DAC locations have higher PLR than those in non-DAC areas.

Recommendation 6:

Consider additional research to determine factors associated with greater in DAC areas, for both third-party operators and homeowners with purchased systems.

NYSERDA Response to Recommendation: Pending. NYSERDA will consider completing additional research to determine which factors are most closely associated with a greater PLR in DAC areas.

2.5.8 Finding 7

Most projects analyzed in this study had five or fewer years of production data. Only a handful had more than seven years. As a result, the performance loss rates developed in this study may not be indicative of loss rates after 10 or more years.

Recommendation 7: Continue to collect production data for the NY-Sun sample included in this study, to establish a larger data set of longer production records that can be used for further study. Explore processes that could be used to collect later production data for the supplemental data sets included.

NYSERDA Response to Recommendation: Pending. NYSERDA will consider this study as part of future performance analyses and/or by studies completed by reputable third parties.

3 Methods

This section summarizes the methods employed to collect production data for sampled sites and analyze production persistence.

3.1 Data collection and review

3.1.1 NY-Sun data

NYSERDA's NY-Sun tracking database in Salesforce provided site-level account information, including installed capacity (kW), application-specific (modeled) production estimations (kWh), Total Solar Resource Fraction (TSRF), array type, system completion date, customer name and contact information, purchase type, installation contractor, and region.

The evaluation contractor team leveraged the Open NY data set for solar projects to identify additional project information, including inverter and panel manufacturer, model, and quantities.[9] Technology type can be identified by cross-referencing manufacturer and model with databases maintained by the California Energy Commission (CEC).[10]

The production data collection effort for this study sought to efficiently coordinate outreach. The objectives of the data collection effort were two-fold:

- Collect production data (in kWh): first-year monthly
- Of lesser priority, the evaluation contractor team sought to complete a short survey with installation contractors to obtain any additional information required to understand the system and production data.

Most of the Large Commercial and Industrial sites with publicly incentivized generation systems provide internet-connected monitoring data to NYSERDA and the public through the DG Integrated Database. Some of the large C&I projects' production data was not available through this resource, in which case site data was collected from the contractor.

[9] Solar Electric Programs Reported by NYSERDA: Beginning 2000, https://data.ny.gov/Energy-Environment/Solar-Electric-Programs-Reported-by-NYSERDA-Beginn/3x8r-34rs

[10] CEC Solar Energy Equipment list: https://www.energy.ca.gov/programs-and-topics/programs/solar-equipment-lists

Residential and small business participating sites' production data was collected through outreach to installation contractors. The evaluation contractor team had contact information through prior data collection efforts. Contractor contact information in the NY-Sun tracking database and the existing contacts were both utilized.

3.1.2 Supplemental residential data

3.1.2.1 Original data provided to the study

The supplemental residential data was used originally to support assessments of project portfolios on behalf of developers or investors. Data provided to this study included monthly production data and project attributes, including project completion date, system size, along with manufacturer and model information. Only New York installed projects, and only project-months that had passed screening checks for the original analysis, were passed on to this study.[11]

For the original analysis, the monthly production data is compared to expected monthly production. High production data is removed from the data set, which is defined as cumulative production for three summer months in excess of 150% of the expected production. Additionally, the first and last months of production are removed from the analysis to avoid incorporating partial months of data.

The capacity factor was not used as a basis for cleaning production data either for its original analysis or for this study. For some projects, the annual capacity factor exceeds reasonable values. It is suspected that the documented system size is based on plans instead of as-built conditions in these cases. If the monthly production is withing the acceptable range of the expected production, these projects are still valid for studying the PLR.

3.1.2.2 Assessment of supplemental data included in the NY-Sun program

The supplemental residential data does not indicate if a project received a NY-Sun incentive. Projects in the supplemental data set were compared to NY-Sun Program tracking data using project location, system size, and completion date. While only a relatively small percentage of projects could be directly identified within the program data, it is still likely that a large

[11] For a detailed comparison of characteristics between NY-Sun sample and NYS supplemental data, please refer to Appendix B.

percentage of the supplemental projects participated in the NY-Sun Program. The majority of projects that did not have a likely match were projects completed in 2016 or earlier and from Long Island.

3.2 Irradiance-normalized annual production data series

3.2.1 Irradiance data cluster analysis

The monthly production data is normalized to calculate the production that a project would have produced in a month with typical or normal solar irradiance for that month of the year. This analysis required normal and actual-year radiation data for the location of each project. For cost and analysis efficiency, projects were clustered into locations with similar irradiation patterns.[12]

The cluster analysis uses a cluster algorithm and visual assessment of the variability in irradiance across the state. The clustering algorithm used groups projects based on climate conditions, geographical proximity, and terrain elevation similarity, ensuring that each cluster is homogenous for these critical factors. The projects within the same cluster use the same actual and Typical Meteorological Year (TMY) weather data for normalizing production. The advantage to this approach is that it limits the amount of irradiance data that is purchased from third-party vendors without compromising the accuracy of the normalization.

To determine the optimal number of clusters, the evaluation contractor team used the "gap statistic" metric as proposed by Tibshirani, Walther, and Hastie (2001), which is a measure of within-cluster variability.[13] The optimal number of clusters was determined to be 29. This was identified by running the clustering algorithm for various numbers of clusters ranging from 10 up to the total number of projects. The gap value variability metric was plotted against the number of clusters. The 29-cluster group was selected because it showed a significant improvement in the gap value compared to the fewer number of clusters. Additionally, the rate of improvement in the gap value decreases beyond 29 clusters yielding diminishing benefits from increasing the number of clusters. The optimal number of clusters are shown in Figure 3-1. The size and color of the

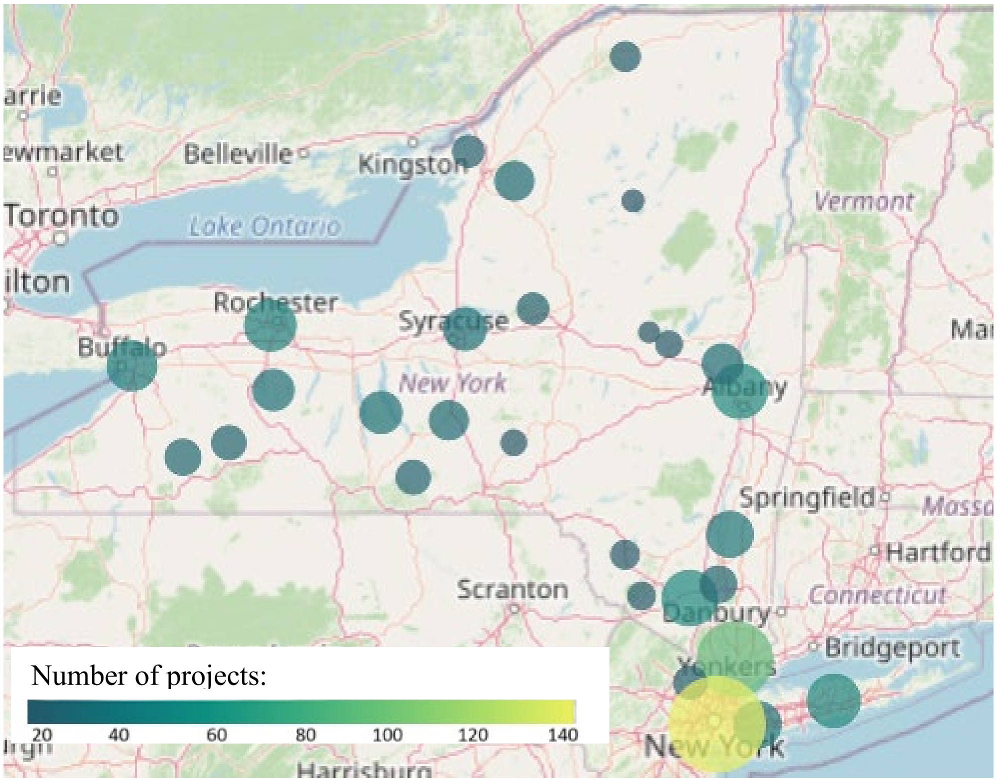
---

[12] The clusters presented in this section correspond to NY-Sun projects. The supplemental data had been previously clustered, and was provided with the irradiance adjustment factors necessary to normalize production.

[13] The gap statistic estimates the number of clusters by comparing within-cluster dispersion to that expected under a reference null distribution.

clusters indicate the number of projects included, with cluster sizes ranging from 2 to 146 projects.

Figure 3-1. 29 clusters selected from optimized cluster analysis



3.2.2 Irradiance normalization

The evaluation contractor team normalized production and capacity factors for weather differences (solar insolation, temperature, etc.) across installation years. The irradiance-normalized values represent performance under typical irradiance conditions and provide a more meaningful basis for comparing changes in production over time. However, this correction does not account for snowfall covering the panels or changes in shading of nearby structures that might affect the solar energy production.

Irradiance adjustments are made with a tool developed by DNV called Sunray. For each month of each year an irradiance adjustment factor is calculated, which indicates the ratio of actual production to the typical year production. The observed production quantity for each site and month is adjusted to TMY conditions by dividing each observed monthly production by the irradiance adjustment.

3.2.3 Data cleaning

For the previous impact analysis studies, the NY-Sun sample data had been cleaned to include only projects with 12 months of production data for the studied year, and excluded projects with unreasonable recorded system size, based on the observed coincidence factors. As indicated in Section 3.1.2, Supplemental residential data, the supplemental data had been screened to exclude projects with summer performance substantially different from its expected production for that period.

The analysis for this study is based on a decay rate in irradiance-normalized monthly production, relative to a basic level that varies by calendar month due to seasonal changes in insolation. Months with 0 or partial production due to temporary disruptions would contribute to large changes down and up from one month to the next, at a level that can mask the more gradual degradation of system performance when operating.

Starting from the previously cleaned NY-Sun sample data used for impact analysis or the cleaned supplemental data used for portfolio assessments, production data was further processed to avoid effects of such disruptions on the performance loss rate analysis, the data was screened by these two criteria:

1. **Disruption exclusion.** For each project, the analysis excluded any month with zero production. Additionally, months that had zero production in the preceding or following month were also excluded. The exclusion of adjacent months means the months with partial production due to components or communications breaking or getting repaired mid-month do not affect the estimated performance loss rate.
2. **Minimum series length**. For monthly analysis using separate PLR estimates for individual projects, only projects with at least 24 months of data were included. This requirement ensures that a meaningful performance loss rate can be estimated for the project, given the variation in underlying production rates due to seasonal patterns.

The results presented in Section 2.3, Results, are based on the analysis of 8,891 projects, following the application of the described cleaning steps.

Table 3-1. Monthly screening attrition

Data set	Number of projects
Initial projects from NY-Sun sample and supplemental data set	13,973
1. Projects in pooled monthly analysis	13,684
2. Projects 2-stage monthly analysis (min 24 months)	**8,981**

Monthly capacity factors: Capacity factor provides a measure of system performance relative to rated capacity. Capacity factor (CF) for a given project is calculated as:

$$CF = \frac{kWh\_eval}{CAP \times days\ in\ Month \times 24\ hrs\ per\ day}$$

Where:

kWh_eval = monthly production for a system (kWh)

CAP = System rated DC capacity

3.3 Performance loss rate analysis

The PLR analysis was conducted on the projects' monthly capacity factors.

3.3.1 Method 1, Project-specific PLR across months

For PLR estimation method 1, the evaluation contractor team estimated a PLR for each project using a linear model, based on the available monthly data for each project. The data for this analysis was limited to projects with at least 24 production months without disruption, as described in Section 3.2, Irradiance-normalized annual production data series. Thie model is a log-linear estimation, where the dependent variable was the logarithm of the capacity factor, and the primary independent variable was the number of months elapsed since the installation date, with monthly fixed effects included. The PLR for each project was estimated using the following equation:

$$\ln\left(CF_{j\,t}\right) = b_{0,j} + \sum_{m=1}^{12} b_{1,m,j}\, M_{mjt} + b_{PLR,j}\, t + \epsilon_{jt}$$

Where:

$\ln\left(CF_{j,t}\right)$ = the natural log of the CF for project j in month t

$b_{0,j}$ = regression intercept

$b_{1,mj}$ = fixed effect for calendar month m

M_{mjt} = 0/1 indicator variable equal to 1 if the observation month t is in calendar month m for project j

t = months elapsed since installation date

$b_{PLR,j}$ = PLR for project j, average monthly production decay rate for project j (percent per month)

The primary results presented in Section 2.3.3, Performance loss rates, are for the project-specific PLR, scaled by 12 to represent the annual average PLR for a given project, averaged across projects for each type within each category of interest. This is a more intuitive framework for considering performance loss rates over time because it corresponds to the average of the individual projects' decay rate by subgroup.

3.3.2 Method 2, Pooled PLR by category, across months

The pooled estimation for Method 2 calculated the PLR by type within a category, across all projects and months in that type. The PLR was calculated using a log-linear regression model that incorporates both fixed effects for each month and interaction variables representing the elapsed time since installation for each category. The model is similar to the project-specific regression, but is estimated across all projects, and allows for a separate PLR for each type with a category (e.g., each region).

$$\ln(CF_{j,t}) = b_j + \sum_{m=1}^{12} b_{1,m} M_{mjt} + \sum_{r} b_{PLR,r}\, Cat_{rj} \times t + \epsilon_{jt}$$

Where:

$\ln(CF_{j,t})$ = the natural log of the CF for project j in observation month t

b_j = fixed effect for project j

$b_{1,m}$ = fixed effect for calendar month m

M_{mjt} = 0/1 indicator variable equal to 1 if the observation month t is in calendar month m for project j

t = months elapsed since installation date

Cat_{rj}= 0/1 indicator variable equal to 1 if project j belongs to category r

$Cat_r \times t$ = this interaction variable represents the product of the months elapsed since the installation date and a category-specific dummy variable. The dummy variable equals 1 if the project category is *r* and 0 otherwise.

$b_{PLR,r}$= PLR for category r, average monthly production decay rate for category r

The results of this estimation, with the category defined as the region, are presented in Table 2-4.

3.3.3 Method 3, Pooled PLR by category and month

The estimation for Method 3 was similar to the across-months regression described in Section 3.3.2, Method 2, Pooled PLR by category across months, but was fit separately for each calendar month. The PLR was calculated by type within a category, across all projects for a particular month of the year. This estimation used a log-linear regression model that included interaction variables to account for the elapsed time since installation for each category, with separate estimates for each month (see equation below).

$$\ln\left(CF_{j,t}\right) = b_j + \sum_r b_{PLR,r}\, Cat_{rj} \times t + \epsilon_{jt}$$

Where:

$\ln\left(CF_{j,t}\right)$ = the natural log of the CF for project j in observation t

b_j = fixed effect for project *j*

t = months elapsed since installation date

Cat_{rj}= 0/1 indicator variable equal to 1 if project j belongs to category r

$Cat_r \times t$ = this interaction variable represents the product of the months elapsed since the installation date and a category-specific dummy variable. The dummy variable equals 1 if the project category is *r* and 0 otherwise.

$b_{PLR,r}$= PLR for category r, average monthly production decay rate for category r

The results of this estimation, with the category defined as the region, are presented in Figure 2-4.[14]

3.3.4 Cross-factor analysis

As noted in Section 2.3.4, Isolating effects of system characteristics on performance loss rates, a cross-factor analysis in the form of a multiple regression analysis was used to attempt to break out the effects of individual project characteristics while controlling for the other characteristics. The analysis was an unweighted categorical regression across all projects.

The evaluation contractor team fitted a regression model of the time coefficient estimated in Section 3.3.1, Project-specific PLR – all months, on project characteristics such as Region and DAC.[15] The regression model was specified as follows:

$$\widehat{b_{PLR,J}} = b0 + b1_{REGj} + b2_{DACj} + b3_{PURCHj} + b4_{MODj} + b5_{INVj} + b7_{INVMj} + e_j$$

Where:

$\widehat{b_{PLR,J}}$= the average monthly decay rate for project j estimated in Section 3.3.3, Pooled regressions by category, individual months

b0 = regression intercept

[14] Coefficients were scaled by a factor of 12 to represent average annual percent PLR.

[15] The regression is limited to the data from the Project-specific PLR, which was limited to projects with at least 24 months of production without disruption as defined in Section 3.2.3, Data cleaning.

$b1_{REGj}$, $b2_{DACj}$, $b3_{PURCHj}$, $b4_{MODj}$, $b5_{INVj}$, $b6_{INVMj}$, respectively, are the coefficients of the region, DAC status, purchase type, module family, inverter type, and inverter manufacturer that project j belongs to.

e_j = residual error of the regression.

While b1-b6 represent the average monthly change in PLR for projects with a given characteristic, the result of this analysis is not intended as the basis for estimating the effects of particular features. Rather, these results confirm and, in some cases, refine qualitative relationships identified from the comparisons of subgroup averages presented in Section 2.3.3, Performance loss rates. The primary results presented in Section 2.3.4, Isolating effects of system characteristics on performance loss rates, are scaled by a factor of 12 to represent the average annual change in PLR for a given characteristic.

Appendix A Persistence study paper

The white paper *DNV'S Views on Long-Term Degradation of PV Systems* is attached on the following pages.

Appendix B Distribution of projects characteristics by data source

Table B-1. Distribution of projects characteristics by data source

Source		NYS supplemental	NY-Sun sample			
Sector		Residential	Residential	Non-residential		
System size (kW)		Below 200 kW	Below 200 kW	Below 200 kW	≥200 to <750 kW	≥ 750 kW
Number of projects		13,587	105	63	139	79
Variable	Category					
Region	Long Island	66%	45%	32%	21%	4%
Region	Downstate	18%	28%	49%	24%	10%
Region	Upstate	17%	28%	19%	55%	86%
DAC	NonDAC	66%	76%	57%	72%	70%
DAC	DAC	34%	24%	43%	28%	30%
Purchase Type	Lease	67%	45%	5%	1%	
Purchase Type	PPA	33%	38%	22%	22%	54%
Purchase Type	Unknown	0%		46%	47%	33%
Purchase Type	Purchase		17%	27%	31%	13%
Module Family	Polycrystalline	16%			1%	1%
Module Family	Unknown	25%	100%	100%	98%	99%
Module Family	Monocrystalline	59%			1%	
Microinverter	String	1%	82%	32%	28%	23%
Microinverter	Optimizer	14%				
Microinverter	Unknown	1%	9%	68%	72%	77%
Microinverter	Microinverter	84%	10%			
Inverter Manufacturer	G	85%				

Source		NYS supplemental	NY-Sun sample			
Sector		Residential	Residential	Non-residential		
System size (kW)		Below 200 kW	Below 200 kW	Below 200 kW	≥200 to <750 kW	≥ 750 kW
Number of projects		13,587	105	63	139	79
Variable	Category					
Inverter Manufacturer	H - Other	1%	5%	33%	25%	23%
Inverter Manufacturer	Unknown		95%	67%	74%	77%
Inverter Manufacturer	I	14%			1%	
Panel Manufacturer	A	12%		3%	1%	8%
Panel Manufacturer	B	36%				4%
Panel Manufacturer	C	3%				
Panel Manufacturer	D	28%				
Panel Manufacturer	E - Other	10%	90%	54%	81%	49%
Panel Manufacturer	F	10%		6%	5%	1%
Panel Manufacturer	Unknown		10%	37%	14%	38%

Appendix C Performance loss rate with region and purchase type distribution

The following appendix contains tables in Section 2.3.3, Performance loss rates, with the region and purchase type distribution of projects in specific subgroups. Table C-1 corresponds to Table 2-6 in the report, and Table C-2 corresponds to Table 2-8 in the report.

Table C-1. PLR by sector and data source (percent production loss per year)

							Regional distribution (percentage of projects)			Purchase type distribution (percentage of projects)			
Category	# of projects	Avg. capacity (kW)	Median # of months	Avg. CF	Mean PLR	SE	Downstate	Island	Upstate	Leased	PPA	Unknown	Purchased
All	8,981	31	59	13%	1.17%	0.10%	19%	55%	26%	82%	15%	2%	1%
Residential	8,704	7	59	13%	1.17%	0.10%	18%	56%	25%	84%	15%	1%	0%
Residential, NY-Sun sample	105	7	84	12%	0.43%	0.29%	28%	45%	28%	45%	38%	0%	17%
Residential, Supplemental NYS projects	8,599	7	59	13%	1.18%	0.11%	18%	56%	25%	85%	15%	1%	0%
Non-residential	277	792	72	12%	1.02%	0.41%	26%	19%	56%	1%	30%	43%	25%
Non-residential, <200 kW	62	103	57	12%	0.45%	0.96%	48%	32%	19%	5%	23%	45%	27%
Non-residential, 200 – 750 kW	136	423	77	12%	1.30%	0.65%	24%	21%	54%	1%	20%	48%	32%
Non-residential, ≥750	79	1970	79	12%	0.98%	0.46%	10%	4%	86%	0%	54%	33%	13%

Table C-2. PLR by subgroup (percent production loss per year) averages of project-specific PLR by subgroup

Category	Type	# of projects	Avg. capacity (kW)	Median # of months	Avg. CF	Mean PLR	SE	Regional distribution (percentage of projects)			Purchase type distribution (percentage of projects)			
								Downstate	Long Island	Upstate	Leased	PPA	Unknown	Purchased
All	All	8981	31	59	12.7%	1.17%	0.10%	19%	55%	26%	82%	15%	1%	2%
Region	Downstate	1670	22	60	12.9%	0.96%	0.23%	100%	0%	0%	84%	11%	2%	3%
Region	Long Island	4946	10	58	13.2%	0.42%	0.13%	0%	100%	0%	79%	19%	1%	1%
Region	Upstate	2365	81	62	12.0%	2.89%	0.21%	0%	0%	100%	85%	10%	1%	3%
Sector	Non-Residential	277	792	72	11.8%	1.02%	0.41%	26%	19%	56%	1%	30%	25%	43%
Sector	Residential	8704	7	59	13.0%	1.18%	0.10%	18%	56%	25%	84%	15%	0%	1%
Size	< 200 kW	8766	7	59	13.0%	1.17%	0.10%	19%	56%	25%	84%	15%	0%	1%
Size	200 - 750 kW	136	423	76.5	11.7%	1.30%	0.65%	24%	21%	54%	1%	20%	32%	48%
Size	> 750 kW	79	1970	79	11.9%	0.98%	0.46%	10%	4%	86%	0%	54%	13%	33%
DAC	NonDAC	6040	32	60	12.8%	1.25%	0.12%	22%	42%	36%	82%	15%	1%	2%
DAC	DAC	2941	29	58	13.1%	1.01%	0.18%	11%	82%	7%	81%	16%	1%	2%
Purchase Type	Lease	7337	7	60	12.8%	0.91%	0.10%	19%	53%	27%	100%	0%	0%	0%
Purchase Type	PPA	1380	82	49	13.2%	2.63%	0.35%	13%	69%	18%	0%	100%	0%	0%
Purchase Type	Purchase	88	380	57.5	12.0%	2.69%	0.74%	30%	31%	40%	0%	0%	100%	0%
Purchase Type	Unknown	176	461	59	11.9%	-0.19%	0.71%	32%	26%	42%	0%	0%	0%	100%
Module Family	Monocrystalline	3031	7	42	13.2%	-0.12%	0.19%	21%	71%	8%	82%	17%	0%	1%
Module Family	Polycrystalline	2131	8	61	13.1%	1.96%	0.19%	18%	49%	33%	80%	19%	0%	0%
Module Family	Unknown	3819	63	60	12.4%	1.76%	0.15%	17%	46%	37%	82%	12%	2%	4%
Microinverter	Microinverter	6750	7	60	13.0%	1.46%	0.11%	19%	55%	25%	86%	14%	0%	0%
Microinverter	Optimizer	1675	7	51	12.7%	0.17%	0.30%	14%	64%	22%	81%	19%	0%	0%
Microinverter	String	272	182	73	12.1%	0.21%	0.50%	27%	32%	42%	46%	16%	7%	31%
Microinverter	Unknown	284	606	67	11.9%	1.00%	0.38%	17%	17%	66%	23%	29%	22%	26%

								Regional distribution (percentage of projects)			Purchase type distribution (percentage of projects)			
Category	**Type**	**# of projects**	**Avg. capacity (kW)**	**Median # of months**	**Avg. CF**	**Mean PLR**	**SE**	**Downstate**	**Long Island**	**Upstate**	**Leased**	**PPA**	**Unknown**	**Purchased**
Inverter Manufacturer	G	6818	7	60	13.0%	1.46%	0.11%	19%	55%	26%	86%	14%	0%	0%
Inverter Manufacturer	H - Other	185	269	68	12.2%	0.03%	0.78%	25%	32%	43%	43%	10%	15%	32%
Inverter Manufacturer	I	1676	8	51	12.7%	0.18%	0.30%	14%	64%	22%	81%	19%	0%	0%
Inverter Manufacturer	Unknown	302	566	78	11.9%	0.90%	0.27%	26%	23%	51%	17%	36%	20%	27%
Panel Manufacturer	A	1637	14	62	12.8%	1.64%	0.11%	18%	47%	35%	98%	2%	0%	0%
Panel Manufacturer	B	1504	10	44.5	13.3%	-0.05%	0.36%	16%	70%	14%	78%	22%	0%	0%
Panel Manufacturer	C	374	7	57	13.3%	2.04%	0.52%	12%	66%	21%	76%	24%	0%	0%
Panel Manufacturer	D	2349	7	55	12.8%	0.77%	0.17%	20%	62%	18%	92%	8%	0%	0%
Panel Manufacturer	Other	1609	80	66	12.3%	1.61%	0.25%	22%	36%	42%	71%	16%	4%	8%
Panel Manufacturer	F	1427	10	61	12.9%	1.89%	0.26%	16%	58%	26%	67%	32%	0%	1%
Panel Manufacturer	Unknown	81	981	65	11.9%	0.37%	0.57%	33%	16%	51%	5%	36%	17%	42%